DIE RECHNERGESTÜTZTE PAARBILDUNG IN KLINISCHEN STUDIEN

Bertram Krauß

1981

Aus dem Institut für Medizinische Informationsverarbeitung, Statistik und Biomathematik der Ludwig-Maximilians-Universität München

Vorstand: Prof. Dr. med. K. Überla

DIE RECHNERGESTÜTZTE PAARBILDUNG IN KLINISCHEN STUDIEN

Inaugural-Dissertation

zur Erlangung der Doktorwürde in der gesamten Medizin an der Ludwig Maximilians-Universität zu München

vorgelegt von

Bertram Krauß

aus

Münchberg

1981

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter: Prof. Dr. rer. biol.hum.
H. K. Selbmann

Mitberichterstatter: Prof. Dr. med. K. Überla

Dekan: Prof. Dr. med. W. Spann

Tag der mündlichen Prüfung: 3. 12. 1981

UNIV.
LEIDEN
BIBL.

DISSERTATIONS - UND FOTODRUCK FRANK GmbH
8000 München 2, Gabelsbergerstraße 15, Tel. 2809090

A b s t r a c t

KRAUSS, BERTRAM:

DIE RECHNERGESTÜTZTE PAARBILDUNG IN KLINISCHEN STUDIEN

Die Arbeit beschreibt die Erstellung eines Programmsystems zur Paarbildung in klinischen Studien und seine Anwendung auf Daten der Münchner Perinatalstudie 1979.

Ausgehend von der Definition zweier unterschiedlicher Vorgehensweisen in klinischen Studien - Studien mit experimentellem und Studien mit nicht experimentellem Ansatz - wird auf das Verfahren des Anpassens und auf die damit verfolgten Ziele eingegangen. Unter Berücksichtigung der bisher bekannten Verfahren wird die Anwendungsproblematik von Anpassungsverfahren aufgezeigt.

Die Beschreibung des realisierten Programmsystems gliedert sich in das eigentliche Anpassungsverfahren, das eine paarweise Zuordnung von Fällen und Kontrollen im Verhältnis 1:1 vornimmt, in ein Analyseprogramm zur statistischen Auswertung der angepaßten Gruppen und einen Kolmogoroff-Smirnoff-Test zur Überprüfung von Häufigkeitsverteilungen zwischen ursprünglicher und angepaßter Gruppe. Dieser Gliederung entsprechend ist es aus drei Programmen aufgebaut, die auf Datenmengen von bis zu 5000 Probanden mit je 25 Variablen angewandt werden können.

Die praktische Anwendung wurde mit einer Stichprobe von 5000 Probanden aus der Münchner Perinatalstudie 1979 durchgeführt. Untersucht wurde, ob sich Auswirkungen der Geburtsüberwachung mit dem Kardiotokogramm auf die Verlegungsrate der Neugeborenen und auf die Häufigkeit der sekundären Sectio ergeben würden. Die Gruppen der Fälle (Geburten mit CTG-Überwachung) und der Kontrollen (ohne CTG-Überwachung) wurden in Bezug auf die qualitativen Variablen "Nationalität", "Familienstand", "Sozialstatus" und "Kliniktyp" und auf die quantitativen Variablen "Alter der Mutter", "Anzahl der Geburtsrisiken" und "Anzahl der befundeten Schwangerschaftsrisiken" angepaßt. Dabei ließen sich keine Einflüsse der CTG-Untersuchung auf die Verlegungsrate und die Häufigkeit der Sectio nachweisen.

Inhaltsverzeichnis

		Seite
1	Grundlagen und Definitionen	1
1.1	Typisierung klinischer Studien	1
1.2	Vergleichbarkeit von Gruppen	3
1.3	Methoden der Anpassung	6
1.3.1	Terminologie	6
1.3.2	"Balancing"	7
1.3.3	"Frequency matching" und "Stratified matching"	8
1.3.4	"Pair matching, individual matching"	9
2	Anwendungsproblematik und -beispiele	9
2.1	Individuelles Anpassen von Paaren	9
2.1.1	Anpassungsverfahren in prospektiven Studien	10
2.1.2	Anpassungsverfahren in retrospektiven Studien	14
2.2	Beispiele zweier Anpassungsverfahren	17
2.2.1	"Guaranteed variable caliper matching"	17
2.2.2	"Minimum distance case-control matching"	20
3	Realisiertes Rechenprogramm zur Paarbildung	23
3.1	Das Anpassungsverfahren	24
3.1.1	Anpassen nach qualitativen Variablen	25
3.1.2	Anpassen nach quantitativen Variablen	28
3.2	Das Analyseprogramm	36
3.3	Der Verteilungs-Test	41

4	Praktischer Anwendungsfall	42
4.1	Das Untersuchungsmaterial	42
4.2	Fragestellungen	44
4.3	Das Anpassungsverfahren	45
4.3.1	Die Anpassungsvariablen	45
4.3.2	Eingabeproblematik	47
4.3.3	Programmdurchführung	52
4.4	Ergebnisse des Anpassungsverfahrens	53
4.5	Zusammenfassung	65
5	Literaturverzeichnis	70

1 Grundlagen und Definitionen

1.1 Typisierung klinischer Studien

Der Begriff der klinischen Studie umfaßt in der Literatur einen weiten Bereich von Termini, die von Medizinern und Epidemiologen in teilweise unterschiedlicher Bedeutung angewandt werden. Häufig vorkommende Begriffe sind z.B. Beobachtungsstudie, Fall-Kontroll-Studie, kontrollierte klinische Studie usw.. Diese Begriffe beschreiben mehr oder weniger exakt die in der Untersuchung angewandten Verfahrensweisen.

Grundsätzlich lassen sich Studien unterteilen in solche mit experimentellem und nicht experimentellem Ansatz. Dies ist eine Unterscheidung nach dem Verhalten des Untersuchers in Bezug auf die zu untersuchende Klientel. In Studien mit experimentellem Ansatz nimmt er aktiv Einfluß auf die Untersuchung und beobachtet die Auswirkungen seines Eingreifens. In der nicht experimentellen Studie, von vielen auch Beobachtungsstudie genannt, verfolgt der Untersucher Entwicklungen bestimmter Populationen ohne aktiv einzugreifen. Der Begriff "Beobachtungsstudie" erscheint mir allerdings unglücklich, weil in jeder wissenschaftlichen Untersuchung beobachtet wird, der Schwerpunkt der Unterscheidung jedoch im aktiven bzw. passiven Verhalten des Untersuchers liegt. Die Untersuchungen mit nicht experimentellem Ansatz lassen sich wiederum aufgliedern in prospektive und retrospektive Untersuchungen. Darunter versteht man unterschiedliche Methoden in der Gewinnung der Daten und in der Beobachtung der Untersuchungspopulation.

In prospektiven Untersuchungen wird das gesamte Vorgehen, das die Auswahl der Population und deren Beobachtungsbedingungen umfaßt, erst exakt geplant und dann nach diesen Richtlinien über einen bestimmten Zeitraum durchgeführt. Die Daten werden in der Reihenfolge ihrer zeitlichen Entstehung gesammelt und genauso wird, ausgehend von möglichen Ursachenfaktoren, ein vermuteter Kausalzusammenhang mit Effekten verfolgt, die im Lauf der Untersuchung entstanden sind. In der prospektiven Untersuchung versucht man also von der Ursache ausgehend auf den Effekt zu schließen.

Retrospektive Untersuchungen verwenden dagegen bereits früher erhobene Datenmengen und versuchen, in rückblickender Weise, Schlüsse aus den vorhandenen Fakten zu ziehen. Dabei vergleicht man Personen, bzw. Personengruppen, die sich in einem bestimmten Effekt, z.B. einer bestimmten Krankheit unterscheiden und bildet eine Fall-Gruppe, die diesen Effekt besitzt, und eine entsprechende Kontroll-Gruppe ohne diesen Effekt. Die gesuchte Antwortgröße ist dann nicht der Effekt, sondern ein möglicher Ursachenfaktor. Es wird also vom Effekt ausgehend auf die Ursache geschlossen. Entsprechend der Benennung beider Vergleichsgruppen wird diese Art der Untersuchung auch als "Fall-Kontroll-Studie" (case-control-study) bezeichnet.

Um die unterschiedlichen Vorgehensweisen von Studien, im Bereich der Datensammlung und bei der Auswahl der zu untersuchenden Probandengruppe, aus der Bezeichnung der Studie exakt entnehmen zu können, prägte Feinstein (9) neue Begriffe.

Er unterscheidet die Erhebung von Daten parallel zum Entstehen von der Erhebung von Daten nach deren Entstehen und bezeichnet dann Studien, die die erstere Methode der Datensammlung verwenden, als "prolective" und solche Studien, die die zweite Methode anwenden, als "retrolective".

Auf die gleiche Weise differenziert er zwei Möglichkeiten die Probanden für eine Studie auszuwählen. Werden sie vor dem Eintreffen eines zu untersuchenden Effektes definiert, dann beschreibt Feinstein diese Studie als "cohort-study". Für eine Studie, die ihre Probanden erst nach dem Eintreten des Effektes definiert, verwendet er die Bezeichnung "trohoc-study", also "cohort" rückwärts gelesen. Prospektive Studien, die ihre Probanden vor dem Eintreffen des zu untersuchenden Effektes definieren und dann die Daten in der Reihenfolge ihres chronologischen Entstehens sammeln, könnte man nach Feinstein als "prolective cohort-study" bezeichnen und retrospektive Studien, die ihre Daten nach dem Entstehen erheben und die Probanden nach dem Eintreffen des Effektes auswählen, entsprechend als "retrolective trohoc-study".

1.2 Vergleichbarkeit von Gruppen

Die bei klinischen Studien erfaßten Probanden werden durch eine Vielzahl von unbekannten und bekannten Stör- und Einflußgrößen bestimmt. Als Beispiel seien nur einige dieser Größen wie Alter, Geschlecht, Beruf, Art

und Schwere der Krankheit, Einzugsgebiet der Klinik, Überweisungspraxis der Ärzte, Verkehrsverhältnisse im Einzugsgebiet, Untersuchung und Datenerhebung bei der Einlieferung usw. genannt.

Um die aus klinischen Studien gewünschten Erkenntnisse über Krankheitsursachen und Behandlungsfolgen zu erhalten, vergleicht man zwei oder mehrere Probandengruppen, wobei, im einfachsten Fall mit zwei Gruppen, die eine Gruppe als Kontrollgruppe dient, an der die Veränderungen der zweiten Gruppe, der Fallgruppe, gemessen wird. Dabei ist es nötig, daß die in der Studie erfaßten Probandengruppen vergleichbar sind, bzw. so ausgewählt werden, daß sie vergleichbar sind und auf diese Weise versucht wird den Fehler, d.h. die Falschinformation der Studie möglichst klein zu halten. Praktisch kommt es dabei auf drei wesentliche Bedingungen an. Repräsentationsgleichheit, Beobachtungsgleichheit und Strukturgleichheit (34).

Repräsentationsgleichheit ist erforderlich, um allgemeingültige Aussagen über die zugrundeliegende Population machen zu können. Die Probandengruppen müssen daher repräsentative Stichproben dieser Population sein. Kann die Repräsentationsgleichheit nicht erfüllt werden, muß die Aussage der Untersuchung entsprechend eingeschränkt werden.

Beobachtungsgleichheit meint die Identität des Beobachtungsverfahrens und der Beobachtungsbedingungen in beiden Gruppen. Sie dient der Vermeidung von "information bias", des systematischen Fehlers in der In-

formation und umfaßt sowohl den Bereich Beobachtung, als auch den Bereich Erfassung der interessierenden Faktoren. Ursachen für solche Informationsfehler wären z.B. der Einfluß des Interviewers auf den Probanden oder die Erhebung von Diagnosen durch verschiedene Untersucher und nach verschiedenen Kriterien.

Strukturgleichheit bedeutet Übereinstimmung der Probandengruppen in den Häufigkeitsverteilungen von Einflußfaktoren, die die Untersuchungsgrößen verändern oder verändern könnten. In Abhängigkeit von der zur Verfügung stehenden Population läßt sich Strukturgleichheit häufig nur durch gezielte Zusammensetzung der Gruppen erreichen.

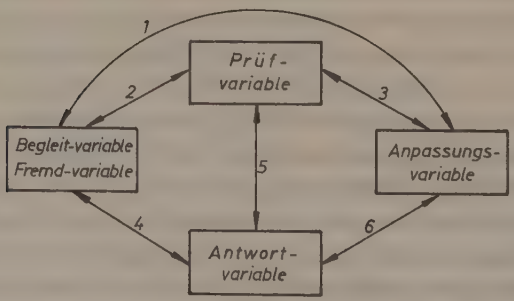
Es sind verschiedene Methoden üblich in diesen drei wichtigen Bereichen Gleichheit zwischen den Probandengruppen herzustellen. Anpassungsverfahren, wie die nachfolgend geschilderten, ermöglichen es nicht, Einfluß auf die Beobachtungsgleichheit zu nehmen. Sie muß durch entsprechende Planung und Durchführung bei der Informationsgewinnung der klinischen Studie gesichert werden (15). Anpassungsverfahren lassen sich auch kaum zur alleinigen Verbesserung der Repräsentationsgleichheit oder der Strukturgleichheit einsetzen, da die durch das Anpassungsverfahren beeinflußten Größen meistens sowohl die Repräsentationsgleichheit als auch die Strukturgleichheit beeinflussen. So sollte das Ziel eines jeden Anpassungsverfahrens sein, die in den oben aufgeführten Gleichheitsbedingungen enthaltenen und teilweise einander entgegenstehenden Prinzipien der Vergleichbarkeit

und der Verallgemeinerungsfähigkeit in einem für die jeweilige Untersuchung optimalen Kompromiß aufeinander abzustimmen.

1.3 Methoden der Anpassung

1.3.1 Terminologie

Zuerst sollen die verwendeten Termini für die relevanten Größen von Anpassungsverfahren erläutert werden. In Anpassungsverfahren kommen vier verschiedene Variablentypen zum Tragen. Die beiden zentralen Variablen sind die P r ü f v a r i a b l e und die A n t w o r t v a r i a b l e (Abb. 1)



2 eliminierbar d. Zufallszuteilung
3 eliminierbar d. Anpassen

Abb. 1: Mögliche Beziehungen und Einflüsse der verschiedenen Variablen

Die Prüfvariable ist diejenige Variable, die im einfachsten Fall zwei Probandengruppen (Fall- und Kontrollgruppe) definiert und die Variable, deren Auswirkung auf die Antwortvariable überprüft werden soll. Die Prüfvariable sollte auf Grund ihrer Funktion dichotom verteilt sein, bzw. dichotom unterteilbar sein. Gewöhnlich beschreibt sie zwei konträre Gruppen wie z.B. Kranke und Gesunde (Prüfvariable: Krankheit) oder unterschiedlich behandelte Probandengruppen (Prüfvariable: Behandlung) wie Patienten mit Therapie 1 und Therapie 2. Die Antwortvariable liefert dem Untersucher die Antwort auf die Bedeutung ihrer Abhängigkeit von der Prüfvariablen. Diese Abhängigkeit bildet die zentrale Fragestellung in einer klinischen Studie mit Anpassen. Die Probanden der Fallgruppe auf der einen und die der Kontrollgruppe auf der anderen Seite sollen bei einer solchen Studie in den Anpassungsvariablen einander angeglichen werden, um theoretisch die Beziehungen der Anpassungsvariablen zur Prüfvariablen (Abb.1: Pfeil 3) auszuschalten. Als Begleitvariable werden die zwar erhobenen jedoch im Anpassungsverfahren nicht verwendeten Variablen und als Fremdvariable die unbekannten Variablen bezeichnet. Ihr Einfluß auf die Prüfvariable läßt sich nur durch zufällige Zuteilung ausschalten (1).

1.3.2 "Balancing"

Die Kontrollgruppe wird bei diesem Verfahren so ausgewählt, daß alle Mittelwerte ihrer Anpassungsvariablen mit den Mittelwerten der Anpassungsvariablen der Fallgruppe übereinstimmen (12).

1.3.3 "Frequency matching" und "Stratified matching"

Bei dieser Methode werden mit den Anpassungsvariablen sogenannte "Matching"-Kategorien gebildet. Dazu wird der Wertebereich einer jeden Anpassungsvariablen in Schichten unterteilt. Eine "Matching"-Kategorie wird gebildet aus der Kombination je einer Schicht aller Anpassungsvariablen. Jeder Kategorie wird bei der Fallgruppe und bei der Kontrollgruppe die gleiche Anzahl von Probanden zugeordnet. Damit wird in beiden Gruppen für jede Anpassungsvariable eine ähnliche Verteilungsfunktion erzielt, was sich mit dem "Balancing" nicht erreichen läßt (2).

"Stratified matching" entspricht dem "Frequency matching". Es werden in gleicher Weise Anpassungskategorien durch Schichtenbildung innerhalb des Wertebereichs jeder Anpassungsvariablen erzeugt und die einander entsprechenden Kategorien der Fall- und Kontrollgruppe in gleicher Zahl besetzt. Cochran (6) beschreibt noch genauer den Besetzungsvorgang der Anpassungskategorien, die er als Zellen bezeichnet. Er zieht aus jeder Zelle der Fallgruppe eine Zufallsstichprobe, deren Größe proportional der ursprünglichen Zellbesetzung ist. Aus den entsprechenden Zellen der Kontrollgruppe werden dann ebenfalls Stichproben gleicher Größe gezogen. Damit wird die Kontrollgruppe in der Verteilung ihrer Anpassungsvariablen der Fallgruppe angeglichen.

1.3.4 "Pair matching", "Individual matching"

Im Gegensatz zu den vorher beschriebenen Verfahren werden beim individuellen Anpassen nicht die beiden Probandengruppen als Ganzes angeglichen, sondern es werden Fall- und Kontrollprobanden zu Einzelpaaren zusammengestellt, die, in dem vom Untersucher vorgegebenen Rahmen, in den interessierenden Variablen übereinstimmen. Dabei ist sowohl die Zuordnung von Probanden der Fall- und Kontrollgruppe im Verhältnis 1:1, als auch in beliebigen Verhältnissen 1:R möglich. Der wichtige Unterschied zu den oben aufgeführten Verfahren besteht darin, daß die beiden Probandengruppen nicht mehr in ihrer Gesamtheit betrachtet werden. Jedes Paar bildet für sich eine Untersuchungseinheit mit eigenen Gesetzmäßigkeiten. Die gesamte Studie setzt sich damit aus sovielen voneinander unabhängigen Einzeluntersuchungen zusammen, wie Paare gebildet worden sind (23).

2 Anwendungsproblematik und -beispiele

2.1 Individuelles Anpassen von Paaren

Das Verfahren des Anpassens wird in klinischen Studien sehr oft anderen Verfahren vorgezogen. Obwohl es ursprünglich als eine "Blockbildungs"-Technik in Experimenten verwendet wurde, in denen ein Paar eine natürliche Blockgröße darstellte, wie z.B. Zwillinge, wurde Anpassen in weiten Bereichen der nicht experimentellen Bevölkerungsstudien als Mittel zur Verminderung der Unterschiedlichkeit von Probandengruppen benützt; nicht

nur in prospektiven, sondern auch in retrospektiven Studien. In zahlreichen Veröffentlichungen wurde das Für und Wider dieser Untersuchungsmethode erörtert. Anhand von Vergleichen mit anderen Verfahren, speziell der Kovarianzanalyse bei zufälliger Stichprobenauswahl, und mittels Modellpopulationen wurden Voraussetzungen erarbeitet, unter denen dieses Verfahren zu einer verbesserten Aussagefähigkeit von Studien führen kann, aber auch aufgezeigt, welche Nachteile daraus erwachsen können. Über bestimmte Probleme, besonders bei der Anwendung in retrospektiven Untersuchungen, herrschen noch differente Meinungen (5,9,10,13,15,20,26,28); die grundsätzlichen Erfordernisse, unter denen individuelles Paaren sinnvoll ist, sind jedoch unumstritten. Trotz dieser Tatsache leidet eine große Zahl medizinischer Untersuchungen, die dieses Verfahren anwenden, an Mängeln, die den Gebrauch dieser Methode als nicht mehr gerechtfertigt erscheinen lassen. Mehrere Autoren haben in Veröffentlichungen auf solche Fehler hingewiesen (2,7,9,13,16,32). Aus diesem Grund möchte ich hier nochmals in einer kurzen Zusammenfassung auf die Probleme dieses Verfahrens in prospektiven und in retrospektiven Untersuchungen eingehen.

2.1.1 Anpassungsverfahren in prospektiven Studien

Prospektive Studien versuchen, durch Beobachtungen in chronologischem Ablauf, Kausalzusammenhänge zwischen bestimmten bekannten Faktoren und ihren vermuteten oder unbekannten Auswirkungen aufzudecken. Als vergleichende Untersuchungen benötigen sie neben der zu

untersuchenden Probandengruppe (Fallgruppe) im einfachsten Fall eine Vergleichsgruppe, die Kontrollgruppe. Beide Gruppen werden, wie eingangs erwähnt, durch die Prüfvariable definiert. Sie ist der Faktor, den z.B. eine der Probandengruppen besitzt, die andere nicht, oder es ist diejenige Exposition, der eine Probandengruppe ausgesetzt ist, die andere nicht. Diese Prüfvariable soll in ihren Auswirkungen auf eine zweite Größe überprüft werden. Die zweite Größe ist die Antwortvariable, deren Veränderungen die Abhängigkeit von der Prüfvariablen beantworten. Effekte, als Ausprägungen der Antwortvariablen, lassen sich dann finden, wenn sichergestellt ist, daß weder die Heterogenität der Beobachtungseinheiten, noch zufällige Einflüsse für diese Effekte allein verantwortlich gemacht werden können. Dabei erfolgt die Ausschaltung von Störfaktoren gewöhnlich durch a) Kovarianzanalyse bei bekannten, quantitativ meßbaren Störgrößen, b) Blockbildung, Schichtung oder durch Paarbildung bei bekannten, nicht meßbaren Störgrößen und c) Randomisierung und Wiederholung bei unbekannten Störgrößen (34). Die hauptsächlichen Beweggründe, die jedoch zur Anwendung des Anpassungsverfahrens führen, scheinen die Übersichtlichkeit der Untersuchungssituation, die vereinfachte Analyse im Vergleich zur Kovarianzanalyse und die vereinfachte Erklärung der Ergebnisse zu sein (2,20).

Das Argument der Übersichtlichkeit muß als schwach bewertet werden. Ein gutes Anpassungsschema zu entwerfen, erfordert genausoviel Information über bedeutende Variable und ihre Interaktion, wie zur Planung einer Datenauswahl im Hinblick auf eine Kovarianzanalyse nötig ist.

Des weiteren bestimmt ein Anpassungsschema die Bedeutung der einzelnen Variablen im voraus, wobei sich aufgrund unserer immer unvollständigen Kenntnis hinterher einige Variable als unnütz erweisen. In der Kovarianzanalyse werden nur die Variablen verwendet, die die Präzision verbessern.

Daß "Pair matching" die statistische Analyse vereinfacht, kann nicht bezweifelt werden; dabei darf jedoch nicht übersehen werden, daß die Sammlung der passenden Stichproben oft schwierig und zeitaufwendig ist. Die damit verbundenen Kostensteigerungen einer Untersuchung haben zu Varianten der Methode geführt, von denen später noch ausführlicher berichtet wird.

Anpassen verbessert ohne Zweifel auch die Genauigkeit von Vergleichen, speziell bei kleinen Untersuchungspopulationen, wo die Zufallsauswahl durch den geringen Umfang oder eine sehr inhomogene Population zu Fall- und Kontrollgruppen sehr unterschiedlicher Zusammensetzung führen kann.

Wie Billewicz (3) mit Hilfe einer, auf einem Computer simulierten, Untersuchung mit einer Modellpopulation von über 14000 fiktiven Probanden zeigte, schneidet die Kovarianzanalyse bei quantitativen Antwortvariablen etwas besser ab, während bei einer dichotom verteilten Antwortvariablen das Verfahren des Anpassens klar überlegen ist und vor der Möglichkeit einer Verzerrung in der Schätzung des Effektes schützt.

Um die Effektivität der verglichenen Verfahren (An-

passen contra Kovarianzanalyse) zu bestimmen, berechnete er die Differenz der Antwortvariablen von Fall- und Kontrollprobanden und die Varianz der Differenz. Ein Verfahren bezeichnete er dann als umso effektiver, je kleiner die Varianz dieser Differenz wurde. Dabei war die Effektivität des Anpassens in dieser Modelluntersuchung abhängig von der Größe der Korrelation zwischen Antwortvariabler und Anpassungsvariablen. Bei höheren Korrelationen zwischen Antwortvariabler und Anpassungsvariablen (Korrelationskoeffizient größer als 0,6) war bei Anwendung des Anpassungsverfahrens die Wahrscheinlichkeit größer einen Behandlungseffekt zu übersehen, als bei Anwendung der Kovarianzanalyse.

Individuelles Anpassen kann darüberhinaus bei der Analyse von prospektiven Untersuchungen Vorteile bieten, wie Smith, Kark, Cassel und Spears (36) gezeigt haben. In dieser Untersuchung bestand ein erheblicher Größenunterschied zwischen der Probandengruppe der Fälle und der Probandengruppe der möglichen Kontrollen (Verhältnis Kontrollen zu Fällen größer als 30:1), wodurch der sonst relevante Nachteil des Verlustes an Probanden durch Anpassen nicht zum Tragen kommen konnte und die Vorteile der Einfachheit und Übersichtlichkeit der Ergebnisse voll zur Geltung gelangen konnten. Ich werde auf diese Methode, da sie einen sehr interessanten Anwendungsaspekt aufzeigt, später noch genauer eingehen.

2.1.2 Anpassungsverfahren in retrospektiven Studien.

Die Anwendung des Anpassungsverfahrens in retrospektiven Untersuchungen muß aus anderen Erwägungen erfolgen. Es ist dabei von Nutzen, die Unterschiedlichkeit beider Untersuchungsverfahren noch einmal deutlich hervorzuheben.
In prospektiven Studien unterscheiden sich die Probandengruppen der Fälle und der Kontrollen durch eine bestimmte Exposition (Prüfvariable). Sie ist der hauptsächliche Unterschied der angepaßten Gruppen und kann dadurch ihren Effekt (Antwortvariable) zeigen, der durch die Kontrolle der ihn beeinflussenden Faktoren (Anpassungsvariable) noch verstärkt werden kann.
In retrospektiven Studien ist es dagegen der Effekt, also z.B. die Krankheit, der den Unterschied zwischen der Probandengruppe der Fälle und der Kontrollen bestimmt. Gemäß der Definition der Prüfvariablen stellt jetzt der Effekt die Prüfvariable dar, während die Exposition die Antwortvariable bildet. Paßt man nun mit Faktoren an, die die Antwortvariable - also in diesem Fall die Exposition - beeinflussen, dann kann die Krankheit, da sie gewöhnlich nicht Teil eines Ursachenkomplexes der Exposition ist, auch keine Veränderungen der Exposition innerhalb der Fall- und Kontrollgruppen hervorrufen, die durch das Anpassen homogen in Bezug auf die wirklichen Determinanten der Exposition sind.
Dies bedeutet, daß bei der prospektiven Studie das Anpassen mit Variablen, die mit der Antwortvariablen korrelieren, die Aussagekraft der Untersuchung erhöht,

während bei der retrospektiven Studie ein Anpassen mit Variablen, die die Antwortvariable beeinflussen, eine Verminderung der Aussagekraft hervorruft. In diesem Fall ist Anpassen nur sinnvoll um die Vergleichbarkeit der Gruppen zu sichern.

Nachfolgend sei kurz die Argumentation Miettinens (26) zusammengefaßt, unter welchen Voraussetzungen es sinnvoll erscheint, eine Variable in einer retrospektiven Studie als Anpassungsvariable zu verwenden und wie sich ihre Einbeziehung in das Anpassungsverfahren auf Validität und Effektivität der Studie auswirkt. Das Ziel der Validität ist dabei das grundsätzlich dominierende. Es beinhaltet die Vermeidung systematischer Fehler bei der Schätzung des untersuchten Effektes durch Verbesserung der Vergleichbarkeit von Fall- und Kontrollgruppe. Die Effektivität bezieht sich auf die Genauigkeit einer Schätzung des Effektes oder auf das Vermögen, die Nullhypothese zu testen, bei einer gegebenen Gesamtzahl von Probanden.

Die Verwendung von angepaßten Kontrollen hat dabei nichts zu tun mit den Eigenschaften der Fälle. Sie beeinflußt nur die Eigenschaften der Kontrollreihe, die dann in Bezug auf die Anpassungsvariablen die gleiche Struktur wie die Fallreihe besitzt. Bedeutend sind die Sekundäreffekte dieser Ähnlichkeit auf die erwartete Verteilung der Antwortvariablen in der Kontrollreihe und die richtige Interpretation der Konsequenzen dieser Effekte.

Besitzt die Variable, mit der angepaßt werden soll, keinen Einfluß auf die Antwortvariable, dann ist eine Auswahl der Kontrollgruppe nach dieser Anpassungsvariablen gleichbedeutend mit einer Wahl, unabhängig von der Fallgruppe, und anschließender zufallsmäßiger Paarung. Unabhängig davon, wie stark die Anpassungsvariable mit der Prüfvariablen korreliert, ist durch Anpassen weder eine Verbesserung noch ein Verlust von Validität und Effektivität möglich, es ist nicht sinnvoll.

Eine Variable, die in Beziehung zur Antwort- und zur Prüfvariablen steht, ist ein klassischer Verschleierungs-Faktor. Ein Anpassen mit ihr sichert die Validität. Eine Effektivitätsbetrachtung ist in diesem Fall problematisch; denn es hat wenig Sinn, die Genauigkeit von Schätzungen ohne und mit Anpassen zu vergleichen, wenn bekannt ist, daß die Schätzung ohne Anpassen einen Fehler unbekannter Größe enthält.

Es bleibt noch die Situation zu betrachten, bei der die Anpassungsvariable zwar mit der Antwortvariablen korreliert, aber nicht mit der Prüfvariablen. Die Anpassungsvariable wird in diesem Fall nicht den Unterschied in den Antwortvariablen bei den Fällen und Kontrollen beeinflussen, vorausgesetzt, sie schließt gleiche Anteile beider Gruppen ein. Anpassen ist zur Erhaltung der Validität nicht nötig und würde die Effektivität verringern, denn zur Schätzung des relativen Risikos würden nur die unterschiedlichen Paare herangezogen, und ihre Zahl

tendiert bei enger Korrelation zwischen Antwort- und Anpassungsvariabler zur Abnahme. Ist man sich über die Verteilung der Anpassungsvariablen bei den Fällen und bei den Kontrollen nicht sicher, bleibt nichts übrig, als doch mit ihr anzupassen oder sie in der Analyse zu kontrollieren.

2.2 Beispiel zweier Anpassungsverfahren.

2.2.1 "Guaranteed variable caliper matching".

Dieses Verfahren wurde 1970 von Althauser und Rubin (1) veröffentlicht. Es wurde in den Jahren 1967/68 entwickelt und versuchte vor allem, zwei Probleme in den Griff zu bekommen: den Verlust von Probanden mit bereits erhobenen Daten durch die Paarbildung ("attrition") und die nicht optimale Anpassung von Paaren ("incomplete matching"), bei denen zwischen den Anpassungsvariablen eines Paares noch Restdifferenzen bestehen bleiben.

Als hauptsächliche Determinanten des Verlustes von Probanden und der nicht optimalen Anpassung werden die Methode der Probandenauswahl (Fall und Kontrolle gleichzeitig oder nacheinander bestimmt), die relative Größe der Populationen, aus denen Fälle und Kontrollen gewählt werden, sowie die Definition und die Bildung eines Paares aufgeführt. Da dieser letzte Punkt allgemein für Anpassungsverfahren relevant ist, soll näher darauf eingegangen werden.

Als wichtige Kriterien zur Paarbildung wurden die folgenden Forderungen aufgestellt:

1. Geringer Abstand ("caliper") , wobei dieses als annehmbar kleine Differenz der Anpassungsvariablen eines Paares definiert ist.
2. Berücksichtigung der Bedeutung der Anpassungsvariablen, mit schmalen erlaubten Anpassungsbereichen für bedeutsame Variable und weiteren Bereichen für weniger wichtige Variable.
3. "Intuitive close matches": Das Anpassungsverfahren sollte garantieren, daß wirklich ähnliche Paare einander zugeordnet werden. Dies betrifft das Problem des Schichtens bei kontinuierlichen Variablen, bei dem Objekte der gleichen Schicht einander unähnlicher sein können als Objekte benachbarter Schichten. Wird als Anpassungsvariable z.B. das Alter der Probanden in Schichten mit einer Abstufung von fünf Jahren eingeteilt (.../31-35/36-40/...), dann paßt zu einem Fall mit dem Alter 35 Jahre eine Kontrolle aus der gleichen Schicht mit 31 Jahren schlechter als eine Kontrolle im Alter von 36 Jahren aus der benachbarten Schicht. Der Anteil solcher Paare, die nicht nahe genug in ihren Variablenwerten beieinander liegen, sollte minimiert werden.

Zu diesen drei, allgemein bedeutsamen Kriterien, ergaben sich weitere Überlegungen zum Problem der Gewinnung möglicher Anpassungspartner. Sie betrafen speziell prospektive Untersuchungen und dienten dazu, den Aufwand für die notwendige Datenerhebung in Grenzen zu halten.

Dazu zählten eine minimale Stichprobengröße (mehr Probanden ergibt höhere Kosten) bei gleichmäßiger Verteilung, d.h. gleicher Anzahl von möglichen Kontrollen für jeden Probanden der Fallgruppe und trotzdem minimale Überlappungen , wobei unter Überlappung die Zuteilung von gleichen Kontrollen auf verschiedene Fälle verstanden wurde.

Das Verfahren, das alle Wünsche der Autoren erfüllte, arbeitet folgendermaßen:
Das Rechenprogramm sucht für jede Anpassungsvariable k und für jeden Proband g aus Gruppe 1 eine Reihe ähnlichster Probanden h aus Gruppe 2. Die gefundene Menge der Probanden pro Proband g und pro k wird als n_k bezeichnet. Die relative Bedeutung der Variablen k wird reflektiert durch die Zahl n_k der ähnlichsten Probanden, dabei ist n_k umgekehrt proportional zur relativen Wertigkeit r_k der Anpassungsvariablen (n_k/r_k = const.). Die Reihen n_k nähester Partner werden dann unterteilt, um zu prüfen, ob schließlich auf jedes g genügend h passen. Wird die Zahl der gewünschten Paarungen (N) beim ersten Suchvorgang nicht für jeden Probanden g erreicht, dann werden nicht mehr ganz so ähnliche Partner hinzugefügt, bis N erreicht ist. Die Anordnung der Anpassungsvariablen nach relativer Bedeutung stellt dabei sicher, daß die größeren Unterschiede hauptsächlich die weniger wichtigen Variablen betreffen. Bei einer ausreichend großen Stichprobe kann die durchschnittliche Bereichsbreite annehmbar klein bleiben, trotz der weiten Bereiche für extremere Paarungen. Vor dem Zuordnen werden noch allen Probanden aus beiden Gruppen

Zufallszahlen zugewiesen, und die zu den Probanden g als passend gefundenen Probanden h werden in der Reihenfolge steigender Zufallszahlen angeordnet. Der Proband h mit der niedrigsten Zufallsziffer wird die erste Paarung für g, der zweite Proband ist sein erster Ersatz usw.

In diesem Verfahren wurde als Besonderheit die variable Bereichsweite angewandt, wobei der Spielraum in der Weite abhängig war von der Bedeutung der Variablen. Damit wurde die intuitiv ähnliche Paarung realisiert und gleichzeitig ein Verlust an möglichen Paaren vermindert. (1)

2.2.2 "Minimum distance case-control matching".

Das zweite zu erwähnende Verfahren wurde von Smith, Kark, Cassel und Spears 1977 (36) veröffentlicht und beschreibt ebenfalls das Vorgehen in einer prospektiven Studie in den Jahren 1960 - 72. Diese als Computerverfahren realisierte Methode wurde zur Analyse speziell der Situation angewandt, in der nur eine relativ kleine Anzahl von Fällen einer untersuchten Krankheit vorkam, während eine große Population möglicher Kontrollen zur Verfügung stand. Die Vorteile der Methode bestehen darin, daß die Ergebnisse leicht zu verstehen und zu interpretieren sind, und daß sich keine Möglichkeit der Entstehung eines systematischen Fehlers durch beeinflussende Variable ergibt, wenn die Paare ähnlich genug gewählt werden. Als Gegensatz dazu wird die multivariate Varianzanalyse

aufgeführt, deren Ergebnisse schwer verständlich und interpretierbar sein können und bei der die Beseitigung systematischer Fehler abhängig ist von den Voraussetzungen des Modells.

In diesem Verfahren werden die anzupassenden Kovariablen nach zwei Typen unterteilt:

a) kategoriale Variable, wie z.B. Geschlecht und Rasse und

b) kontinuierliche Variable wie Alter und Gewicht.

Bei den kategorialen Variablen wird auf einer perfekten Anpassung, d.h. Identität der Variablenwerte bestanden. Für kontinuierliche Veriable wird ein Abstandsmaß benötigt, um die Ähnlichkeit zweier Probanden zu bestimmen. Die Verfasser verwenden hier einen Summenabstand über alle Anpassungsvariablen. Die einzelnen Variablenwerte werden dazu auf z-Normalwerte transformiert unter Verwendung der Mittelwerte und der Standartabweichungen der Anpassungsvariablen der Fälle. Der Abstand D zwischen dem Fall und einem möglichen Kontrollprobanden ist die Summe der quadratischen Abstände der einzelnen Variablen, die zu dem behandelten Paar gehören ($D = \sum (z_{mi} - z'_{mi})^2$). z_{mi} ist der normierte Wert für die i-te Variable m_i des Falles, z'_{mi} ist der entsprechende Wert für die mögliche Kontrolle. Der Grund für die Normierung der Variablenwerte liegt in der angestrebten gleichen Gewichtung aller einbezogenen Anpassungsvariablen. Der Untersucher kann die Gewichtung variieren, falls er dies vorzieht. Zur Berechnung der Normwerte werden deshalb die Mittelwerte und Standartabweichungen allein der Fälle herangezogen, um keine Veränderungen der z-Werte zu erhalten, falls eine andere Reihe von Kontrollen verwendet wird.

Das Rechenprogramm paßt sequentiell jedem Fallprobanden eine Kontrolle an, indem es das ähnlichste Individuum auswählt, das nicht schon vorher zugeteilt wurde. Bei gleicher Eignung werden die möglichen Kontrollen zufallsmäßig ausgewählt.
Zur Auswertung wurden Formeln für den X^2 - Test und das relative Risiko nach Mantel und Haenszel (23) verwendet. Die Vertrauensbereiche wurden nach einer Methode von Miettinen (27, 29) ermittelt. Darüberhinaus wurden Möglichkeiten aufgezeigt, Interaktionen der Anpassungsvariablen mit den Auswirkungen der Risikofaktoren auf die Antwortvariable darzustellen.

Damit zeigten die Verfasser auf, daß die von vielen Autoren vertretene Ansicht nicht haltbar ist, daß der Gebrauch von Schichten in der Analyse, möglicherweise in Verbindung mit Kovarianzmethoden, grundsätzlich überlegen ist. Ihre Methode umgeht die Nachteile des Schichtens, die in der begrenzten Zahl von verwendbaren Variablen und im Verlust an Information bei Anwendung auf kontinuierliche Variable liegen, und die unter den anfangs gemachten Voraussetzungen der geringen Fallzahl besonders schwer wiegen würden.

3 Realisiertes Rechenprogramm zur Paarbildung

Das Konzept für das realisierte Rechenprogramm ergab sich aus den zu lösenden Aufgaben.
Diese Aufgaben setzten sich zusammen aus der Bildung von Paaren aus vorgegebenen Datenbeständen, aus der statistischen Analyse des Anpassungsverfahrens und aus der Möglichkeit, die beiden Probandengruppen der Fälle und Kontrollen in Bezug auf die Verteilung einzelner Variablen überprüfen zu können.
Dabei sollte es ferner möglich sein, zu einem späteren Zeitpunkt Programmteile auszutauschen oder Ergänzungen einzufügen.
Aus diesen Anforderungen ergab sich eine Aufteilung des Gesamtkomplexes in einzelne Abschnitte mit den oben aufgeführten Funktionen:

1. Das eigentliche Anpassungsverfahren mit einer paarweisen Zuordnung von Probanden der Fall- und Kontrollgruppe im Verhältnis 1 : 1.
2. Ein Analyseprogramm, das die einander angepaßten Gruppen vergleicht und statistisch auswertet.
3. Eine Testroutine, um eventuelle Verteilungs-Änderungen zwischen angepaßter und ursprünglicher Gruppe erkennen zu können. (Abb. 2)

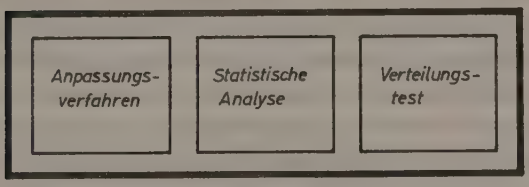


Abb. 2: Programmsystem "Pair matching"

3.1 Das Anpassungsverfahren.

Aus der Größe vorliegender Datenbestände ergab sich der Wunsch, möglichst umfangreiche Probandengruppen erfassen zu können. Dieser Wunsch wird eingeschränkt durch zwei einander bedingende, jedoch entgegengerichtete Fakten: die Kapazität des Arbeitsspeichers der Rechenanlage auf der einen Seite, der Zeitaufwand zum Lauf des Programms auf der anderen Seite. Umfangreiche Datenblöcke müssen zur Bearbeitung von externen Massenspeichern wie Magnetplatte oder Band gelesen und danach auf diese wieder ausgelesen werden. Diese Lese- und Schreibvorgänge erfordern im Vergleich zur Rechenzeit einen relativ hohen Zeitaufwand. Die Größe der maximal zu bearbeitenden Datenmenge wurde deshalb so gewählt, daß bei maximal ausgenütztem Speicherplatz möglichst wenig Lese- und Schreibvorgänge nötig wurden. In der jetzt realisierten Version können Dateien mit bis zu 5000 Probanden und pro Proband bis zu 25 Variablen bearbeitet werden. Bei einer Datei dieser Größe und einem Anpassen auf 7 Variable benötigt das gesamte Programmpaket mit Auswertung nach 2 Antwortvariablen und einer Verteilungsüberprüfung einer Variablen eine Rechenzeit von ca. 500 sec.

Funktionell arbeitet das Anpassungsverfahren in zwei Schritten. Dies wird bedingt durch den Typus möglicher Daten, die die Eigenschaften der Probanden beschreiben. Diese Daten lassen sich grundsätzlich in zwei verschiedene Typen unterteilen, in qualitative und in quantitative Variable. Dabei beschreiben quantitative Variable Eigenschaften ihrer Träger, die sich in Meß- oder

Zählgrößen angeben lassen. Als Beispiel seien hier Körpermaß oder Zeit als stetige Merkmale und Kinderzahl als diskretes Merkmal aufgeführt.
Im Gegensatz dazu lassen sich qualitative Variable nicht mengen- oder größenmäßig erfassen. Zu ihnen gehören Alternativmerkmale (ja-nein, Geschlecht), kategorial gegliederte Merkmale (wertungslose Reihenfolgen wie z.B.: Augenfarbe, Berufe) und ordinale Merkmale (natürliche Reihenfolgen wie Rangfolgen, Schmerzintensität usw.). Die beiden Anpassungsschritte werden in zwei Programmteilen verwirklicht (Abb.3), deren erster eine Aufschlüsselung aller Probanden nach Fall- und Kontrollgruppe und nach Identitätsklassen auf Grund qualitativer Variabler umfaßt. Der zweite Teil ordnet dann innherhalb der entstandenen Klassen die Probanden einander zu, die durch Definition des Untersuchers nach quantitativen Variablen ähnlich sind.

3.1.1 Anpassen nach qualitativen Variablen.

Zu diesem ersten Programmschritt muß der Anwender die Prüfvariable und ihre zugehörigen Merkmalswerte sowie die qualitativen Anpassungsvariablen und deren zugehörige Qualitäten definieren und eingeben. Die Prüfvariable muß dabei, allein schon durch ihre Funktion als Differenzierungsmerkmal von Fall- und Kontrollgruppe, von vornherein dichotom verteilt oder entsprechend unterteilbar sein.
Mit den angegebenen Informationen werden die möglichen Kombinationen der Qualitätswerte aller Anpassungs-

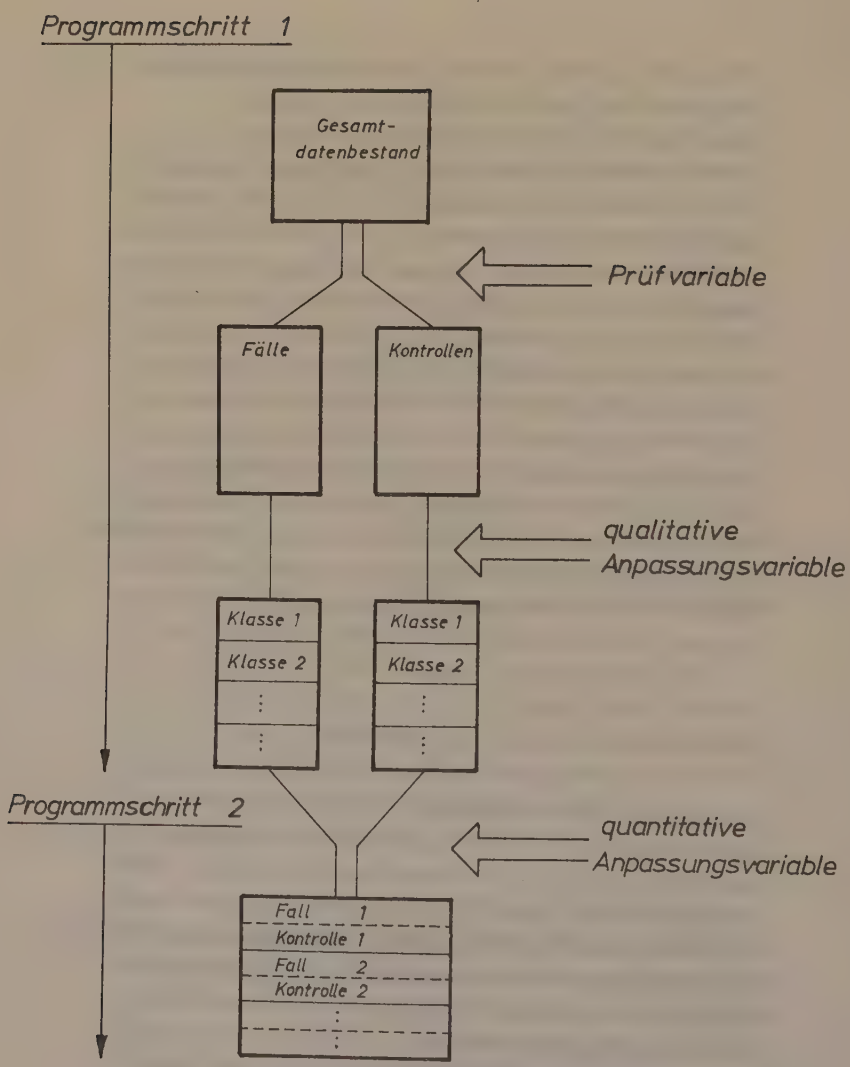


Abb. 3: Funktionelle Gliederung des Anpassungs-
verfahrens.

variablen ermittelt und als Klassen gekennzeichnet, getrennt nach Fall- und Kontrollgruppe. Die Datensätze, von denen jeder einen Probanden mit allen Variablen repräsentiert, werden auf Grund ihrer Qualitätswerte klassifiziert und erhalten Zufallsziffern zugeteilt. In den nachfolgenden Schritten werden für beide Gruppen die klassifizierten Datensätze nach identischen Klassen und innerhalb jeder Klasse nach aufsteigenden Zufallsziffern geordnet.

Dieser erste Abschnitt des Anpassungsverfahrens entspricht in seiner Vorgehensweise dem Schichten, ohne jedoch eine Anpassung im eigentlichen Sinn vorzunehmen. Diese würde ja in allen Klassen Verteilungshomogenität durch gleiche Besetzungsziffern herstellen. Die Beschränkung auf die Behandlung der qualitativen Variablen umgeht die Nachteile, die Schichten besitzt: den möglichen Informationsverlust durch eine Klassifizierung von kontinuierlichen Variablen, wobei die Klassengrenzen einander beliebig ähnliche Objekte unterschiedlichen Gruppen zuteilen und die hohe Zahl an Schichten, die bei diskreten Variablen entstehen kann und von der Graduierung ihrer Meßskalen abhängt.

Die Besetzungsziffern jeder Klasse, die ermittelt werden, liefern dem Untersucher eine Vorinformation über Verteilungsunterschiede zwischen Fall- und Kontrollgruppe und ermöglichen damit Aussagen über Zusammenhänge von Anpassungsvariablen mit der Prüfvariablen und der Anpassungsvariablen untereinander. Unnütze Anpassungsvariablen können damit erkannt und aus dem Verfahren genommen werden.

Obwohl konzipiert für eine Untersuchung mit bereits erhobenen Daten, läßt sich diese Vorstufe des "Sortierens" auch in prospektiven Studien benützen, um Informationen über noch fehlende Kontrollgruppen zu erhalten und um die Gewinnung von Daten entsprechend auslegen zu können.

Die in diesem Abschnitt ermittelten Größen werden für die folgenden Programmteile zwischengespeichert. Der Kontrollausdruck für den Anwender enthält die Prüfvariable mit der Definition des Wertebereichs für die Fall- und Kontrollgruppe, eine Liste der verwendeten qualitativen Variablen und ihrer Qualitäten, eine Erläuterung, welche Klasse durch welche Qualitätswerte definiert ist und die Verteilung der Gesamtpopulation, dargestellt durch Besetzungsziffern der einzelnen Klassen von Fall- und Kontrollgruppe.

3.1.2 Anpassen nach quantitativen Variablen.

Dieser zweite Programmabschnitt führt das eigentliche paarweise Anpassen mit quantitativen Variablen durch. Das Verhältnis der Zuordnung von Kontrollprobanden zu Fallprobanden ist hier 1:1 gewählt. Die Aufteilung des Anpassungsverfahrens in einzelne Funktionsabschnitte ermöglicht es jedoch, an dieser Stelle beliebige Varianten der Anpassung, wie z.B. Mehrfachzuordnungen von Kontrollen, als Alternativen einzufügen.

Das "Sortieren" im ersten Programmschritt beschränkt nun die aufwendigere Anpassung nach quantitativen Variablen auf die Datenmenge jeweils einer Klasse und verringert auf diese Weise den Zeitaufwand für die Suche nach ähnlichen Probanden.
Quantitative Variable ließen sich zwar prinzipiell gleich behandeln wie qualitative Variable; doch die Unterteilung ihrer Werteskalen wirft die bereits im vorhergehenden Abschnitt angesprochenen Probleme der Bereichswahl bei kontinuierlichen Variablen, der Zugehörigkeit zu einem Bereich und die mögliche hohe Zahl an Schichten bei kleinen Bereichsbreiten auf. Aus diesen Gründen wurde eine Möglichkeit, quantitative als qualitative Variable zu behandeln, also eine Anpassung durch "stratification" vorzunehmen, nicht vorgesehen.

Ähnlichkeit bei quantitativen Variablen muß nicht Identität bedeuten, wie dies bei qualitativen Variablen der Fall ist. Der Untersucher definiert hier einen ausreichend kleinen Abstand, innerhalb dessen zwei Probanden einander ähnlich sind. Er bestimmt über den Abstand zugleich auch eine Güte der Anpassung, wenn er je nach Bedeutung, die er einer Variablen beimißt, deren zulässige Abstandsdifferenz gering oder weit wählt. Offen bleibt damit, was Ähnlichkeit über die Summe der zu berücksichtigenden quantitativen Anpassungsvariablen bedeutet, da Ähnlichkeit in Bezug auf eine Variable die komplexen Beziehungen und Einflüsse aller Anpassungsvariablen nicht erfassen kann. Weiterhin muß entschieden werden, wie unter mehreren möglichen Paarungen, die bei Erfüllung der Ähnlichkeitskriterien als gleichwertig zu betrachten sind, ausgewählt wird.

Die Abstandswahl kann nach verschiedenen Gesichtspunkten getroffen werden. Eine Möglichkeit ist, nur den gemittelten Abstand über alle Variablen zu betrachten, wie es die Autoren Smith, Kark, Cassel und Spears (36) in ihrer Arbeit beschreiben. Die Vorteile liegen in einer gleichmäßigen, bzw. durch den Untersucher bestimmbaren Gewichtung aller Variablen und in einem Ausgleich von extremen Variablenwerten durch andere, sehr nahe Werte, womit die Verluste an Kontrollprobanden verringert werden. Dieser Ausgleich kann jedoch auch ein Nachteil sein, da man ohne nachgeschaltete Tests nie weiß, wie eng die Fall- und Kontrollgruppen in Bezug auf einzelne Variable angepaßt sind.
Eine zweite Methode ordnet jeder einzelnen Anpassungsvariablen einen Mindestabstand zu, innerhalb dessen jeder Kontrollproband ähnlich in Bezug auf diese Variable ist. Man erhält für jeden Probanden der Fallgruppe und für jede Anpassungsvariable eine Reihe möglicher Kontrollen. Die Reihenfolge der Anpassungsvariablen, nach der dann die Paarzuordnung erfolgt, wäre gleichbedeutend mit einer Rangfolge der Variablen. Die erste zugeteilte Variable würde automatisch die höchste Wertigkeit, d.h. Gewichtung erhalten, da sie die Auswahlmöglichkeiten bei den folgenden Variablen einschränkt. Daraus folgt, daß der Untersucher eine Reihenfolge der Anpassungsvariablen in Abhängigkeit ihrer Bedeutung festlegen muß. Dies ist das Vorgehen wie es Althauser und Rubin (1) erwähnen. Diese beiden Methoden ließen sich noch miteinander kombinieren und würden damit den Nachteil der ersten, daß zu extreme Variablenwerte bei Anpassungen einzelner Variablen Ver-

wendung finden, und den der zweiten, daß die Ähnlichkeit von einer Reihenfolge und damit von einer Bewertung der Anpassungsvariablen abhängt, einschränken.

Neben der Definition der Ähnlichkeit durch einen über den gesamten Wertebereich einer quantitativen Variablen gleichbleibenden Abstand ist auch eine prozentuale Ähnlichkeitsdefinition denkbar, die die zulässige Weite eines Abstandes in Abhängigkeit vom Wert der Anpassungsvariablen festlegt. Für diese Methode ergeben sich die gleichen Konsequenzen in Bezug auf die Gewichtung der Variablen wie bei der Definition eines gleichbleibenden Abstandes. Schließlich bleibt die Entscheidung zu treffen, welche der möglichen Kontrollen dem Fallprobanden zugeordnet werden soll, wenn ähnliche Probanden der Kontrollgruppe auf die gleichen oder einige der gleichen Probanden der Fallgruppe passen. die Größe dieser "Überlappung" ist abhängig von der Definition der Ähnlichkeit. Läßt man breitere Anpassungsbereiche zu, ist die Zahl der zueinander passenden Probanden in beiden Gruppen entsprechend größer, als bei sehr eng gewählten Bereichen. Eine Reduzierung der Überlappung durch "exaktere" Anpassung geht dann auf Kosten der Probanden, deren Variablenwerte extremer liegen und die naturgemäß weniger korrespondierende Partner in der anderen Gruppe besitzen. Ein einzugehender Kompromiß wird daher immer in die Richtung einer Erhöhung der Paarzuordnung tendieren und lieber eine Überlappung in Kauf nehmen. Eine Verteilung der entstandenen Mehrfachanpassungen kann dann entweder zufallsbedingt oder über aufwendigere Methoden, mit dem Ziel einer maximalen Anzahl von

Paaren durchgeführt werden. Algorithmen für maximale Zuordnungen sind aus der Graphentheorie bekannt (8). Eine Berücksichtigung dieser Methoden im Anpassungsverfahren wäre jedoch, bezogen auf den erreichbaren Effekt, zu aufwendig.

Realisiert wurde nun folgende Methode: Ausgangspunkt ist die, im ersten Programmschritt nach Fall- und Kontrollgruppe aufgeteilte, klassifizierte und nach steigenden Zufallszahlen geordnete Probandenpopulation. Die Ähnlichkeit zweier Probanden wird über zulässige Maximalabstände jeder einzelnen Anpassungsvariablen definiert. Für jeden Probanden der Fallgruppe wird die entsprechende Klasse der Kontrollgruppe durchsucht, bis ein Proband gefunden ist, der in allen quantitativen Variablen innerhalb der vorgegebenen Grenzen übereinstimmt. Die erste zufallsmäßig gefundene Paarung wird akzeptiert.

Mit diesem Verfahren sind auf einfache Art alle Bedingungen einer gültigen Anpassung gelöst, die auch von den vorher besprochenen diffizileren Methoden erreicht wurden:

1. Die Vermeidung von unkontrolliert großen Abständen bei einzelnen Anpassungsvariablen durch Festlegung eines maximal zulässigen Abstandes.
2. Die Möglichkeit, die Bedeutung von Anpassungsvariablen durch die Breite des Anpassungsbereiches zu berücksichtigen.

   Die Anordnung aller Probanden nach steigenden Zufallsziffern beseitigt zudem

3. eine unbeabsichtigte Gewichtung der Anpassungsvariablen durch ihre Reihenfolge und
4. das Problem der Überlappungen.

Die gepaarten Probanden scheiden aus dem weiteren Verfahren aus. Die Gesamtpopulation wird so Klasse für Klasse auf Paarungen durchsucht und alle einander zugeordneten Probanden paarweise gespeichert.

Auf einem Kontrollausdruck erhält der Anwender, wie im ersten Programmschritt, Angaben über die quantitativen Variablen und die gewählten Ähnlichkeitsbereiche und eine, wiederum nach Klassen geordnete, Liste über die Zahl der Fälle der zugehörigen Kontrollen und der erzielten Paarungen. Der gesamte Datenbestand wird inclusive aller bisher ausgewerteten Größen zwischengespeichert und steht den folgenden Programmteilen oder zum Ausdrucken zur Verfügung.
Das Anpassungsverfahren läßt die zur Eingabe benötigte Datei unverändert und erzeugt aus den übernommenen Daten drei Ergebnis-Dateien.
Die erste Ergebnisdatei ist vom Inhalt her mit der Eingabedatei identisch, enthält also alle Probanden und unterscheidet sich nur in deren Anordnung.
Sie enthält der Reihe nach:

1. Alle Fälle nach qualitativen Klassen geordnet; innerhalb jeder Klasse geordnet nach gepaarten Fällen und ungepaarten Fällen. Die gepaarten Fälle stehen immer am Anfang jeder Klasse. Ihre korrespondierenden Partner bei den Kontrollen stehen in der gleichen Klasse auf gleicher Position.
2. Alle Kontrollen nach dem gleichen Schema geordnet wie die Fälle.

3. Alle Fälle, die durch fehlende Angaben oder fehlerhafte Daten nicht zum Anpassen herangezogen wurden.
4. Alle Kontrollen mit fehlenden Angaben oder fehlerhaften Daten.
5. Alle Probanden, die durch fehlerhafte oder nicht definierte Daten weder der Gruppe der Fälle noch der Gruppe der Kontrollen zugeordnet werden konnten.

Mit Hilfe dieser Datei lassen sich Verteilungsunterschiede bei den gepaarten und ungepaarten Fall- und Kontrollgruppen ermitteln.
Die Ergebnisdateien 2 und 3 enthalten nur die gepaarten Fälle und Kontrollen. Dabei ist die Datei 2 so geordnet, daß sie erst alle gepaarten Fälle und dann alle gepaarten Kontrollen - geordnet nach aufsteigenden Klassen - enthält. Fall und zugehörige Kontrolle stehen innerhalb ihrer Klasse auf der gleichen Position. In der Datei 3 stehen dagegen Fall und zugehörige Kontrolle unmittelbar hintereinander (Abb. 4).

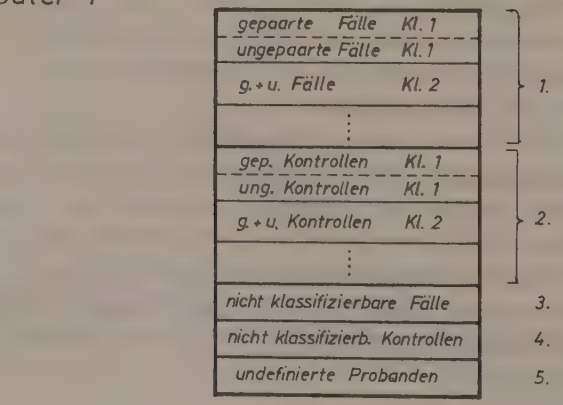
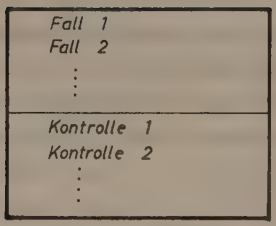


Abb. 4.1: Struktur der Ergebnisdatei 1

Datei 2



Datei 3

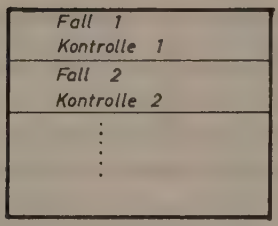


Abb. 4.2: Struktur der Ergebnisdateien 2 und 3

3.2 Das Analyseprogramm.

Die Ergebnisse von Anpassungsverfahren lassen sich, unter der Voraussetzung von dichotom verteilten Prüfvariablen und ebenfalls dichotom verteilten Antwortvariablen, in der üblichen 4-Felder-Tafel darstellen.

	AV(+)	AV(-)	Σ
PV(+)	A	B	N_1
PV(-)	C	D	N_2
Σ	M_1	M_2	T

Ein allgemein gebräuchlicher Test auf Assoziation beider Merkmale ist der X^2 - Test, anhand dessen die Nullhypothese auf Unabhängigkeit beider Merkmalsalternativen abgelehnt oder beibehalten wird. Eine korrigierte Formel für die Prüfgröße wurde von Yates vorgeschlagen:

$$\hat{X}^2 = \frac{(|A \cdot D - B \cdot C| - {}^1/_2\, T)^2}{N_1 \cdot M_1 \cdot N_2 \cdot M_2} \qquad (34)$$

Werden Zufallsstichproben von Individuen mit und ohne Prüfvariabler in Klassen verschiedener Kontrollfaktoren (Anpassungsvariable) eingeteilt, dann kann die Verteilung der Antwortvariablen innerhalb der i-ten Klasse als Vierfeldertafel dargestellt werden.

	$AV_i(+)$	$AV_i(-)$	Σ
$PV_i(+)$	A_i	B_i	N_{1i}
$PV_i(-)$	C_i	D_i	N_{2i}
Σ	M_{1i}	M_{2i}	T_i

Die korrigierte Formel für $\hat{\chi}^2$ ergibt sich aus

$$\hat{\chi}^2 = \frac{(|A_i - E(A_i)| - \frac{1}{2})^2}{V(A_i)} \qquad (23)$$

d.h. der Unterschied zwischen A_i und dem Erwartungswert von A_i, unter der Hypothese eines einheitlichen relativen Risikos ($E(A_i) = N_{1i}M_{1i}/T_i$), wird relativ zur Varianz $V(A_i)$ getestet, die, vorbehaltlich fester Randsummen, durch $V(A_i) = N_{1i}N_{2i}M_{1i}M_{2i}/\ T_i^2(T_i-1)$ gegeben ist. Diese Formel für die Varianz wird gebildet aus der Varianz der Binomial-Variablen N_1PQ ($P=M_1\ /\ T$; $Q=M_2/T$) multipliziert mit einem Korrekturfaktor für begrenzten Stichprobenumfang $(T-N_1)/(T-1) = N_2/(T-1)$. Die Formel nach Yates gebraucht dagegen einen Korrekturfaktor von N_2/T. Der Unterschied

zwischen beiden Gleichungen besteht in einem Faktor (T-1)/T, der zwar für große T zu vernachlässigen ist, jedoch bei einem Signifikanztest über alle Klassen Bedeutung erlangen kann. Besteht zwischen der Prüf- und Antwortvariablen eine stetige Beziehung in der gleichen Richtung über die verschiedenen Klassen, dann ergibt sich die Formel für die Prüfgröße $\hat{\chi}^2$:

$$\hat{\chi}^2 = \frac{(|\Sigma A_i - \Sigma E(A_i)| - \frac{1}{2})^2}{\Sigma V(A_i)}$$

mit $E(A_i)$ und $V(A_i)$ wie oben angegeben.

Eine Studie mit individuellem Anpassen kann man als Sonderfall einer Klassifikations-Prozedur betrachten, bei dem die Anzahl der Klassen gleich der Anzahl der individuellen Paare ist. Trägt man den Status eines jeden Paares, bestehend aus je einem Individuum mit PV(+) Fallproband und PV(-) Kontrollproband, gegen die An- bzw. Abwesenheit der Antwortvariablen auf, erhält man die folgende Vierfelder-Tafel:

		PV(+)		
		AV(+)	AV(-)	Σ
PV(-)	AV(+)	F	G	C
	AV(-)	H	I	D
Σ		A	B	N

F, G, H und I entsprechen der Anzahl der Paare mit der jeweiligen Kombination von Prüf- und Antwortvariablen. Die Randsummen entsprechen den vorhin aufgeführten Größen.

Für F der N Klassifikationen (= Paare) gilt dann:

$A_i=1;\ M_{1i}=2;\ M_{2i}=0;\ E(A_i)=1;\ V(A_i)=0,$

für G: $A_i=0;\ M_{1i}=1;\ M_{2i}=1;\ E(A_i)=1/2;\ V(A_i)=1/4,$

für H: $A_i=1;\ M_{1i}=1;\ M_{2i}=1;\ E(A_i)=1/2;\ V(A_i)=1/4,$

für I: $A_i=0;\ M_{1i}=0;\ M_{2i}=2;\ E(A_i)=0;\ V(A_i)=0.$

Damit ergibt sich $A_i = F+H;\ E(A_i) = F+ 1/2(G+H);$

$V(A_i) = 1/4(G+H)$ und damit:

$$\hat{\chi}^2 = \frac{(|G-H|-1)^2}{(G+H)}$$

Das gleiche Ergebnis erhält man auch durch eine andere Überlegung. Besteht zwischen PV(+) und AV(+) keine Beziehung, dann erwarten wir die gleiche Anzahl von AV(+) sowohl bei PV(+) als auch PV(-), d.h. wir erwarten, daß A (= F+H) = C (= F+G). Dies tritt nur dann ein, wenn G = H ist. Der Test vergleicht die Häufigkeiten von G und H und prüft, ob sie eine deutliche Abweichung vom Verhältnis 1 : 1 aufweisen, d.h. ob sich G signifikant von 1/2(G+H) unterscheidet. G wird als Binomialvariable mit dem Parameter 1/2 getestet, G+H ist die Gesamthäufigkeit, G hat den Erwartungswert 1/2(G+H) und die Varianz 1/4(G+H).

Damit wird: $\hat{\chi}^2 = \frac{(|G-H|-1)^2}{(G+H)}$

Dies ist die korrekte Form des Tests von individuell angepaßten Paaren (23). Bei vielen Untersuchungen wird jedoch die ganz am Anfang besprochene Form des χ^2 - Tests angewandt. Dies wäre nur dann korrekt, wenn die Wahrscheinlichkeit des Vorhandenseins der Antwortvariablen (AV(+)) für alle Paarungskategorien als gleich angenommen werden könnte, d.h. wenn AV(+) unabhängig wäre von allen Anpassungsvariablen. In diesem Fall gäbe es eigentlich keinen Grund für ein Anpassen, es wäre irrelevant.

Im verwirklichten Programm werden beide Vierfeldertafeln ermittelt. Die erste kann bei Anpassungsverfahren verwendet werden, bei denen die Identität der Paare nicht beibehalten wird, wie z.B. beim "frequency matching", und benötigt genügend hohe Besetzungszahlen in jeder Kategorie und damit insgesamt große Stichproben, auch wenn nur wenige Anpassungskategorien benützt werden.

Im Ergebnisausdruck zu diesem Programm wird nochmals auf die unterschiedliche Definition der zwei Vierfelder-Tafeln eingegangen und die kritischen Werte für Signifikanzniveaus von 0,1%, 1% und 5% der χ^2 - Verteilung für den zweiseitigen Test angegeben. Die ermittelten Feldgrößen und die zugehörigen χ^2 - Werte für beide Felder werden im Anschluß daran aus-

gedruckt. Zusätzlich für Feld 2 (für individuelles paarweises Anpassen) wird nach Formel von Mantel und Haenszel (23) das relative Risiko R = H/G und die "Odds ratio" R1 = AD/BC berechnet. Es folgt noch eine Aufschlüsselung der einzelnen Feldgrößen nach den im qualitativen Anpassen festgelegten Klassen. Weitergehende Angaben zur Analyse von Studien mit Anpassen sind der Literatur zu entnehmen. (23, 27, 29, 30, 32).

3.3 Der Verteilungs-Test.

Als Option für den Anwender wurde noch der Kolmogoroff-Smirnoff-Test in das Programmsystem eingefügt, um eventuelle Verteilungsveränderungen von quantitativen Variablen durch den Anpassungsvorgang erfassen zu können. Es werden, für eine vom Untersucher zu definierende quantitative Variable, bei der gepaarten und der ungepaarten Fall - und Kontrollgruppe die absoluten und die relativen Summenhäufigkeitskurven ermittelt und ausgedruckt. Als Prüfgröße dient das Maximum der Absolutbeträge der relativen Differenzen:

$$D_{max} = max \left| \frac{F_1}{n_1} - \frac{F_2}{n_2} \right|$$

Der kritische Wert, bei dessen Überschreitung ein bedeutsamer Verteilungsunterschied vorliegt, wird für eine Irrtumswahrscheinlichkeit von 1% ermittelt (34).

4. Praktischer Anwendungsfall.

4.1 Das Untersuchungsmaterial.

Das realisierte Anpassungsverfahren wurde auf Datenmaterial angewandt, das im Rahmen der Münchner Perinatalstudie 1979 gesammelt wurde. Aus einem Gesamtdatenbestand von 54 686 Einlingsgeburten in Schädellage wurden über Angaben für die Indikation zur operativen Entbindung alle Entbindungen mit primärer Sectio aussortiert. Folgende Indikationen wurden dazu herangezogen:

1. Plazenta praevia
2. vorzeitige Plazentalösung
3. Gestose - Eklampsie
4. Zustand n. Sectio
5. Zustand n. anderen Uterusoperationen
6. absolutes oder relatives Mißverhältnis zwischen kindlichem Kopf und mütterlichem Becken
7. Plazentainsuffiziens
8. Mehrlingsschwangerschaften

Aus dem verbleibenden Datenbestand wurde eine Zufallsstichprobe von 5000 Probanden gezogen. Die Daten eines jeden Probanden umfaßten 21 Variable. Zusätzlich wurde jedem Probanden eine Identifikationsziffer zugeteilt. Einen Überblick über die in der Datei gespeicherten Variablen gibt Tabelle 1.

Tabelle 1

Qualitative und quantitative Variable der Untersuchungspopulation.

Qualitative Variable:

- Mortalität
- Verlegung in die Kinderklinik
- Nationalität
- Entbindungsart
- Familienstand
- Sozialstatus
- CTG ante partum
- CTG sub partu intern
- CTG sub partu extern
- Haltungs- und Einstellungsanomalien
- Kliniktyp

Quantitative Variable:

- Parität
- Geburtsgewicht
- Apgar nach 1 min.
- Alter der Mutter
- Anzahl der anamnest. Schwangerschaftsrisiken
- Anzahl der befundeten Schwangerschaftsrisiken
- Anzahl der Geburtsrisiken
- Tragzeit
- Anzahl der Vorsorgeuntersuchungen

4.2 Fragestellungen.

In Bezug auf das vorliegende Datenmaterial interessierte, ob sich Auswirkungen der Geburtsüberwachung mit einem Kardiotokogramm (CTG sub partu) bei der Verlegungsrate der Neugeborenen als grobem Mortalitätsmaß und bei der sekundären Sectio feststellen ließen. Wie eingangs bereits erläutert, steht in dem vorliegenden Datenmaterial der Begriff Sectio grundsätzlich für sekundäre Sectio, da alle primären Sectiones ausgeschlossen wurden (vgl. 35).
Im Gegensatz zu üblichen retrospektiven Untersuchungen, die rückblickend von einer Wirkung auf die Ursache zu schließen versuchen, soll hier die Untersuchung ausgehend von der Ursache (CTG s.p.) in Richtung möglicher oder fraglicher Wirkungen (Verlegungsrate, sek. Sectio) erfolgen. Unter Verwendung der von Feinstein (9) vorgeschlagenen Begriffe also eine "retrolectiv cohort"- Studie.

Auf der anderen Seite sollten erste Erfahrungen im Umgang mit den erstellten Programmen gemacht werden, die sich einmal auf den praktischen Teil, die Handhabung der einzelnen Programmteile erstrecken, zum anderen auf den funktionellen Bereich. Hier interessierten vor allem die Auswirkungen, die die Programmstruktur auf die Ergebnisse hat und ob von den Programmen ausreichende Information geliefert wird, um eine Transparenz des Anpassungsvorganges zu ermöglichen. Weiterhin sollten die Kontrollmöglichkeiten in Form des Kolmogoroff-Smirnoff-

Testprogramms und das Auswerteprogramm zum Anpassungsverfahren überprüft werden.

4.3 Das Anpassungsverfahren.

4.3.1 Die Anpassungsvariablen.

Die Auswahl der geeigneten Anpassungsvariablen erfordert bei einem guten Anpassungsverfahren genügend Kenntnisse über die einzelnen Variablen und ihre Beziehungen zueinander. In diesem Fall konnte auf die durchgeführten Analysen der Münchner Perinatal-Studie 1975-1977 (35) zurückgegriffen werden, um die Bedingungen zu prüfen, denen eine Anpassungsvariable genügen sollte. Die Variablen sollten sowohl die Prüfvariable (CTG s.p.) als auch die Antwortvariable (Verlegungsrate, Sectio) beeinflussen. Wird ein solcher Einfluß nur vermutet, sollten auch solche Variable zum Anpassen herangezogen werden. Als qualitative Variable wurden in das Verfahren die in Tabelle 2 aufgeführten Variablen aufgenommen. Die Variablen Sozialstatus, Nationalität und Familienstand lassen sich als Faktoren sozialer Gruppen betrachten. Diese Gruppen unterscheiden sich hinsichtlich ihres Anteils an der Untersuchungspopulation, der Schwangerschaftsüberwachung und der Geburtsergebnisse (siehe 35 S. 64ff, S. 75-77), d.h. die die Gruppen bildenden Variablen beeinflussen unsere Prüf- und Antwortvariablen. Ebenso zeigen die Kliniktypen Unterschiede hinsichtlich der CTG-Überwachung, des Schwangerenkollektivs, der Sectio und der Verlegungsrate (35, S. 52ff.).

Tabelle 2

Qualitative Anpassungsvariable:

Nationalität	0 = deutsch	
	1 = ausländisch	
Familienstand	0 = alleinstehend	
	1 = verheiratet	
Sozialstatus	0 = keine Angabe	
	1 = niederster	
	2 = niederer	
	3 = gehobener	
	4 = höherer	
Kliniktyp	0 = Universitätsklinik	
	1 = Chefarztklinik	≤ 1000 Betten
	2 = Chefarztklinik	> 1000 Betten
	3 = Belegarztklinik	≤ 300 Betten
	4 = Belegarztklinik	> 300 Betten

Als quantitative Variable wurden das Alter der Mutter, die Anzahl der Geburtsrisiken und die Anzahl der befundeten Schwangerschaftsrisiken zur Anpassung verwendet. Die berücksichtigten Geburtsrisiken sind aus Tabelle 3, die Schwangerschaftsrisiken aus Tabelle 4 zu entnehmen.

Der Einfluß der Variablen Geburtsgewicht auf das Verfahren wird in einem ersten Schritt untersucht. Dazu wird das Anpassungsverfahren mit einer jeweils veränderten Bewertung der Variablen Geburtsgewicht durchgeführt und die Ergebnisse verglichen. Die Bewertung erfolgt durch die Wahl des erlaubten Anpassungsbereichs, wobei der Bereich $\pm\infty$ der geringsten Gewichtung und der Bereich $\pm$ 100 g der stärksten Gewichtung entspricht.

4.3.2 Eingabeproblematik.

Die Eingabe der benötigten Parameter erfolgt bei allen Programmen im Dialogverfahren, wobei schrittweise die einzugebenden Daten angefordert werden. Sind dabei jedoch eine größere Anzahl von Variablen einzugeben und soll das Programm eventuell mehrere Male mit gleichen oder mit leicht veränderten Eingangsgrößen angewandt werden, dann beansprucht die Dialogeingabe zu viel Zeit, so daß es sich als günstiger erwies, die gesamten Eingabedaten in einer Datei zu erstellen und diese Datei dem Programm als Eingabemedium zuzuweisen. So lassen sich auch Eingabewerte einfach verändern, ohne daß die unverändert gebliebenen Parameter nochmals eingegeben

Tabelle 3

Liste der berücksichtigten Geburtsrisiken.

- Grünes Fruchtwasser
- Frühgeburt (Gewicht des Kindes unter 2500g, bzw. Geburt vor Ende der 37. Schwangerschaftswoche)
- Geburtsdauer über 12 Stunden
- Fieber unter der Geburt (>38,0°C rektal/subl.)
- Blutungen unter der Geburt
- Querlage
- Beckenendlage
- Vorderhauptlage
- Gesichtslage
- Tiefer Querstand
- Hoher Gradstand
- Nabelschnurvorfall
- Mehrlingsgeburt
- absolutes oder relatives Mißverhältnis zwischen kindlichem Kopf und mütterlichem Becken
- Acidose sub partu (pH< 7,20)
- Blutdruckanstieg s.p. (RR syst. >140, diast. >90)
- Sonstige

Tabelle 4

Befundete Schwangerschaftsrisiken

EPH-Gestose (Blutdruck über 140/90, Eiweißausscheidung 1‰ oder mehr, mittelgradige Ödeme oder Gewichtszunahme pro Woche von 500g oder mehr)

Harnwegsinfekt, auch Bakteriurie mit Keimzahlen über 100 000/ml

Diabetes mellitus - internistisch behandlungsbedürftig

Blutungen in der Schwangerschaft

Blutgruppen-Inkompatibilität

Mißverhältnis zwischen Uterus- bzw. Kindsgröße und Schwangerschaftsdauer (z.B. Riesenkind, retardiertes Kindswachstum, Hydramnion)

Vorzeitige Wehentätigkeit (vor Ende 37. Woche)

Cervixinsuffizienz

Mehrlinge

pathologische Kindslage

Hypotonie im 3. Trimenon (RR syst. < 100)

Anämie in graviditate (Hb unter 11,2g% gleich 70%)

Überschreitung des Geburtstermins (mehr als 9 Tage über ET)

Unklarheit über den Termin

Sonstige befundete Risiken

werden müssen.

Dabei stellt sich auch die Frage, wie der Wertebereich der Prüfvariablen möglichst anpassungsfähig definiert werden kann. Die Prüfvariable ist im Programm ja diejenige Variable, die die Fälle und die Kontrollen definiert. Dazu sollte sie idealerweise dichotom verteilt sein, bzw. mindestens als dichotom verteilt behandelt werden können. Um dies für verschiedene Typen von Variablen mit unterschiedlichen Codierungen verwirklichen zu können, wurden zwei Bereiche eingeführt. PV(+) bezeichnet den Bereich zur Definition der Fälle und PV(-) den Bereich zur Definition der Kontrollen. Der Anwender hat die Möglichkeit, die für ihn zweckmäßigste Definition zu wählen, wobei er entweder PV(+) oder PV(-) oder beide zusammen zur Auswahl hat. Die Definition erfolgt durch die Angabe einer oberen und einer unteren Grenze des Wertebereichs der Prüfvariablen, der den entsprechenden Bereich PV(+) oder PV(-) darstellen soll. Damit lassen sich sowohl qualitative wie quantitative Prüfvariable verwenden. Wird nur der Bereich PV(+), also der Fall definiert, dann wird der Bereich PV(-) durch den gesamten zu PV(+) komplementären Wertebereich der Prüfvariablen gebildet und umgekehrt. Durch die Möglichkeit,beide Bereiche definieren zu können, lassen sich bei qualitativen Variablen nicht verwendbare Qualitätswerte und entsprechend bei quantitativen Variablen nicht benötigte Wertebereiche aus dem Verfahren ausklammern. Zur Verdeutlichung sei als Beispiel die Variable "CTG ante partum" aus Tabelle 1 aufgeführt.
Würde man als "Fall" alle Probanden mit dem Qualitätswert "2"(CTG = ja) definieren, dann bestünden die

Kontrollen aus den Probanden mit den Qualitätswerten "1" (CTG = nein) und "0" (CTG = keine Angabe). Die Probanden mit fehlender Angabe könnte man durch die zusätzliche Definition des Bereichs PV(-) = "1" ausschließen. In gleicher Weise wie hier für eine Prüfvariable geschildert, stellt sich das Problem auch bei der Antwortvariablen, mit deren Hilfe die möglichen Auswirkungen oder Folgen der Prüfvariablen gemessen werden sollen.

Bei einer umfassenden Verschlüsselung von qualitativen Variablen werden Qualitätswerte vorkommen, die fehlende Angaben oder fehlende Beurteilungsmöglichkeiten verschlüsseln. Solche Qualitätswerte lassen sich verständlicherweise nicht zum Vergleich auf Ähnlichkeit heranziehen. Sie müssen aus dem Anpassungsverfahren ausgeschlossen werden. Da alle möglichen Verschlüsselungen vorkommen können, muß der Anwender des Programms die in seinem Datenmaterial auftretenden und zum Anpassen benötigten Qualitätswerte aller qualitativen Variablen definieren. Probanden, bei denen das Programm nicht definierte Qualitätswerte erkennt, werden nicht im Anpassungsverfahren berücksichtigt. Auf diese Weise wurden in unserem Anwendungsfall alle Probanden mit fehlenden Angaben über den Sozialstatus aussortiert.

4.3.3 Programmdurchführung.

Wie vorhin bereits kurz erwähnt, sollte zuerst der Einfluß des Geburtsgewichts untersucht werden und - in Abhängigkeit vom Ergebnis - dann diese Variable in die weiteren Untersuchungen einbezogen oder ausgeschlossen werden. Dazu wurde nach dem Anpassen nach qualitativen Variablen eine mehrfache Anpassung nach den quantitativen Variablen "Alter der Mutter", "Anzahl der Geburtsrisiken", "Anzahl der befundeten Schwangerschaftsrisiken" und "Geburtsgewicht" durchgeführt, wobei die letzte Variable jeweils mit verändertem Ähnlichkeitsbereich ε eingesetzt wurde. Nach jedem Anpassungsprozeß wurden die Summenhäufigkeitskurven des Geburtsgewichts bei Fällen und Kontrollen aufgestellt und die Abhängigkeit der Verlegungsrate und der Sectio von der Kardiotokographie-Untersuchung durch den X^2 - Wert ermittelt. Dieser Untersuchungsabschnitt wurde mit einer reproduzierbaren Zufallszuteilung von Fällen und Kontrollen vorgenommen, d.h. die Anordnung der Fälle und Kontrollen erfolgte immer nach der gleichen Zufallszahlenreihe, so daß unterschiedliche Paarzuordnungen bei den wiederholten Anpassungsprozessen nicht auf dem Einfluß der Zufallszuteilung beruhen konnten, sondern durch die veränderte Gewichtung der Variablen Geburtsgewicht verursacht sein mußten.

Im nächsten Erprobungsschritt wurde bei konstanter Variablendefinition die Zufallszuteilung nicht mehr reproduzierbar gewählt, so daß die entstehenden Paarungen von den veränderten Zufallsanordnungen

der Fälle und Kontrollen abhingen. Die Summenhäufigkeitskurven des Geburtsgewichts und das Auftreten der Verlegung und der Sectio bei den Paaren wurde wie oben ermittelt. Auf die Ergebnisse und ihre Interpretation soll im folgenden Abschnitt eingegangen werden.

4.4 Ergebnisse des Anpassungsverfahrens.

Der erste Programmteil, das Anpassen nach qualitativen Variablen, ist, wie anfangs geschildert, nicht als Anpassen im eigentlichen Sinn zu verstehen, sondern als Sortieren nach Fällen und Kontrollen, Klassifizieren nach den Werten der qualitativen Variablen und Aussondern von Probanden, die sich nicht als Fall oder Kontrolle einordnen oder auf Grund nicht definierter Variablenwerte nicht klassifizieren lassen. Demzufolge liefert dieser Abschnitt dem Anwender Information über die Zahl der Fälle und Kontrollen, über die Zahl der ausgesonderten Probanden und über die Verteilung der Fälle und Kontrollen auf die einzelnen Klassen. Der Begriff "Klasse" beschreibt eine der möglichen Kombinationen aus den Werten der qualitativen Variablen. In unserem Anwendungsbeispiel ergaben sich aus den Variablen Nationalität und Familienstand mit je zwei Qualitäten, Sozialstatus mit vier Qualitäten und Kliniktyp mit fünf Qualitäten insgesamt 80 Klassen, von denen jedoch nicht alle besetzt waren.
Zahlenmäßig ergab sich eine Aufteilung von 4091 Fällen, d.h. Geburten mit kardiotokographischer

Überwachung, und 909 Kontrollen. Damit wurden fast 82% der in der Stichprobe erfaßten Geburten kardiotokographisch überwacht. Durch fehlende Angaben bei der Variablen "Sozialstatus" wurden bei den Fällen 92, bei den Kontrollen 32 Probanden aussortiert, so daß zur Paarbildung 3999 Fälle 877 Kontrollen gegenüberstanden. Die Zahl der maximal erzielbaren Paare wird natürlich durch die geringe Zahl an Kontrollen bestimmt. Welche Auswirkungen diese unterschiedliche Besetzung der Fall- und Kontrollgruppe hervorruft wird später noch behandelt.

Der Kontrollausdruck mit den Besetzungsziffern der einzelnen Klassen ermöglicht z.B. die Aufstellung von Häufigkeiten der qualitativen Variablen in Abhängigkeit von der Prüfvariablen (Tabelle 5) und bietet damit die Möglichkeit, Aussagen über die Bedeutung der einzelnen Variablen zum Anpassungsverfahren und ihre Beziehungen zueinander und zur Prüfvariablen zu gewinnen.

Während beim Anpassen nach qualitativen Variablen die Probanden einer Klasse in den qualitativen Variablenwerten identisch sind, bleibt es beim Anpassen nach quantitativen Variablen dem Untersucher überlassen, die Ähnlichkeit der Probanden durch die Angabe einer maximal erlaubten Wertedifferenz zu bestimmen. Diese Ähnlichkeit ε wurde bei den quantitativen Variablen "Alter der Mutter" auf $\varepsilon = \pm 3$ Jahre, "Anzahl der Geburtsrisiken" auf $\varepsilon = \pm 0$ und "Anzahl der befundeten Schwangerschaftsrisiken" auf $\varepsilon = \pm 1$ festgelegt und bei allen Programmläufen unverändert beibehalten.

Tabelle 5

	Nationalität deutsch	ausländ.	
Fälle	3580	419	3999
Kontr.	750	127	877
	4330	546	4876

	Familienstand alleinst.	verheir.	
Fälle	307	3692	3999
Kontr.	79	798	877
	386	4490	4876

	Sozialstatus 1	2	3	4	
Fälle	40	879	2494	586	3999
Kontr.	16	226	507	128	877
	56	1105	3001	714	4876

	Kliniktyp 0	1	2	3	4	
Fälle	363	1258	814	513	1051	3999
Kontr.	14	196	227	144	296	877
	377	1454	1041	657	1347	4876

Durch die variable Gewichtung des "Geburtsgewichts" mit $\varepsilon_1 = \pm\infty$; $\varepsilon_2 = \pm$ 500 g, $\varepsilon_3 = \pm$ 250g, $\varepsilon_4 = \pm$ 100g ergaben sich nach dem Anpassen die in der Tabelle 6 aufgeführten Ergebnisse.

Die strenger werdende Forderung in Bezug auf Übereinstimmung nach dem Geburtsgewicht läßt verständlicherweise die Zahl der gefundenen Paare kleiner werden. Eine Auswirkung der Variablen "Geburtsgewicht" auf die untersuchten Abhängigkeiten von "Verlegungsrate" und "sek.Sectio" von der Variablen "CTG" ist nicht vorhanden, wie sich aus den nur gering differierenden X^2 - Werten für beide Antwortvariablen ablesen läßt.

Untersucht man die Abhängigkeit des Faktors "Geburtsgewicht" von der Prüfvariablen "CTG s.p." und stellt zu diesem Zweck die relativen Summenhäufigkeitskurven für Fälle und Kontrollen vor dem Anpassen auf, erhält man das in Abb. 5 dargestellte Ergebnis. Man erkennt, daß zwar die Summenkurve der Kontrollen im unteren Gewichtsbereich geringfügig über der Kurve der Fälle liegt, jedoch keine auffälligen Unterschiede erreicht. Dies bedeutet, daß bereits vor dem Anpassen hinsichtlich der Variablen "Geburtsgewicht" keine auffälligen Verteilungsunterschiede bei den Fällen und Kontrollen bestehen. Auf der anderen Seite ist das Geburtsgewicht jedoch mit der Antwortvariablen "Verlegungsrate" und "sek. Sectio" korreliert (35 S.91, 125, 129).
Wie im theoretischen Teil bereits erläutert, ist eine Anpassungsvariable, die zwar mit der Antwortvariablen,

Tabelle 6

Ergebnisse des Anpassungsverfahrens bei unterschiedlicher Gewichtung der Variablen "Geburtsgewicht".

$\varepsilon_{gewicht}$	Paare	Antwortvariable Verlegungsrate G	H	χ^2	sek. Sectio G	H	χ^2
± ∞	832	45	43	0,011	28	18	1,761
± 500 g	799	36	35	0,000	26	16	1,929
± 250 g	760	33	35	0,015	23	15	1,289
± 100 g	693	27	21	0,521	23	13	2,250

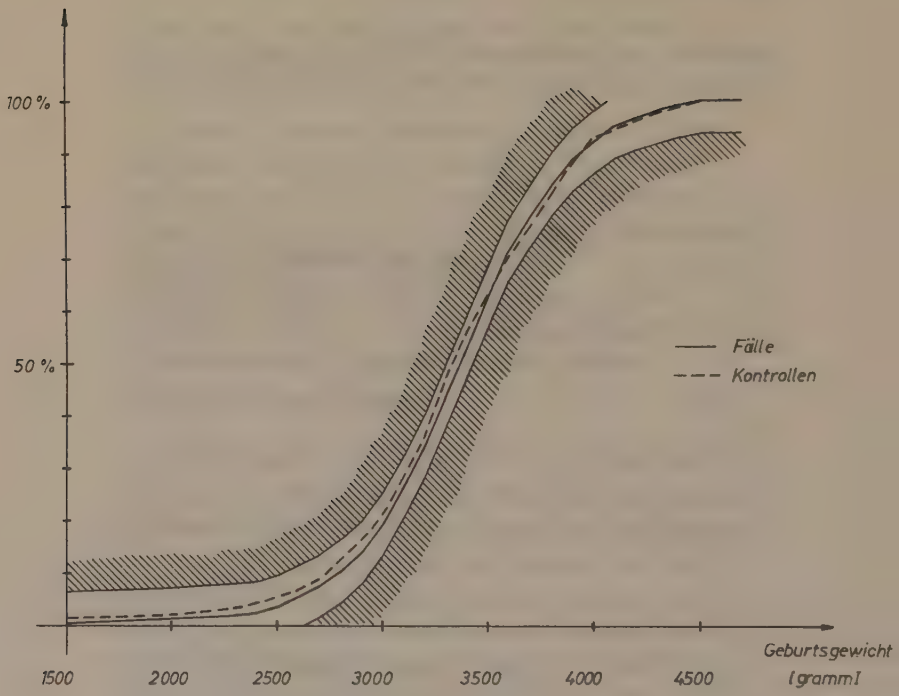


Abb. 5: Relative Summenhäufigkeiten der ungepaarten Fälle und Kontrollen.
Es ist der Bereich gekennzeichnet, innerhalb dessen Grenzen (5% Signifikanzniveau) der Unterschied zwischen Fällen und Kontrollen unauffällig ist.

jedoch nicht mit der Prüfvariablen korreliert, zur Erhaltung der Validität nicht nötig. Sie verringert nur die Effektivität durch die Verringerung der Zahl der unterschiedlichen Paare (G, H), wie aus der Tabelle 6 zu entnehmen ist. (Der Wert G gibt die Anzahl der Paare an, bei denen der Fall den Antwortbereich "negativ" (AV(-)), die Kontrolle den Bereich "positiv" (AV(+)) besitzt. Der Wert H beschreibt die Anzahl der Paare mit umgekehrter Zuordnung, also Fall mit AV(+) und Kontrolle mit AV(-). Aus den Werten G und H errechnet sich die Prüfgröße nach

$$\hat{X}^2 = \frac{(|G-H| - 1)^2}{(G+H)}).$$

Die quantitative Variable "Geburtsgewicht" wurde deshalb bei den folgenden Untersuchungen über die Auswirkungen der Zufallsteilung nicht mit in das Anpassungsverfahren einbezogen. Der Kolmogoroff-Smirnoff-Test ergab bei all diesen Paarzuordnungen auch weiterhin nach dem Anpassen keine bedeutsamen Unterschiede im Geburtsgewicht bei Fällen und Kontrollen.

Die Ergebnisse, die die Anpassungsverfahren mit variabler Zufallszuteilung lieferten, sind in Abb. 6 zusammengefaßt. Es sind die Größen G und H für die Antwortvariablen "Verlegungsrate" und "sek.Sectio" in einem Diagramm aufgetragen. Die eingezeichneten Parabeläste stellen die Funktion $\hat{X}^2_1 = 3{,}841$ (Signifikanzniveau von 5%) und $\hat{X}^2_2 = 6{,}635$ (Signifikanzniveau von 1%) dar. Dabei bedeutet ein Punkt innerhalb der Parabeläste ein nicht signifikantes Ergebnis ($\hat{X}^2 < \hat{X}^2_{1,2}$), während ein Punkt auf oder

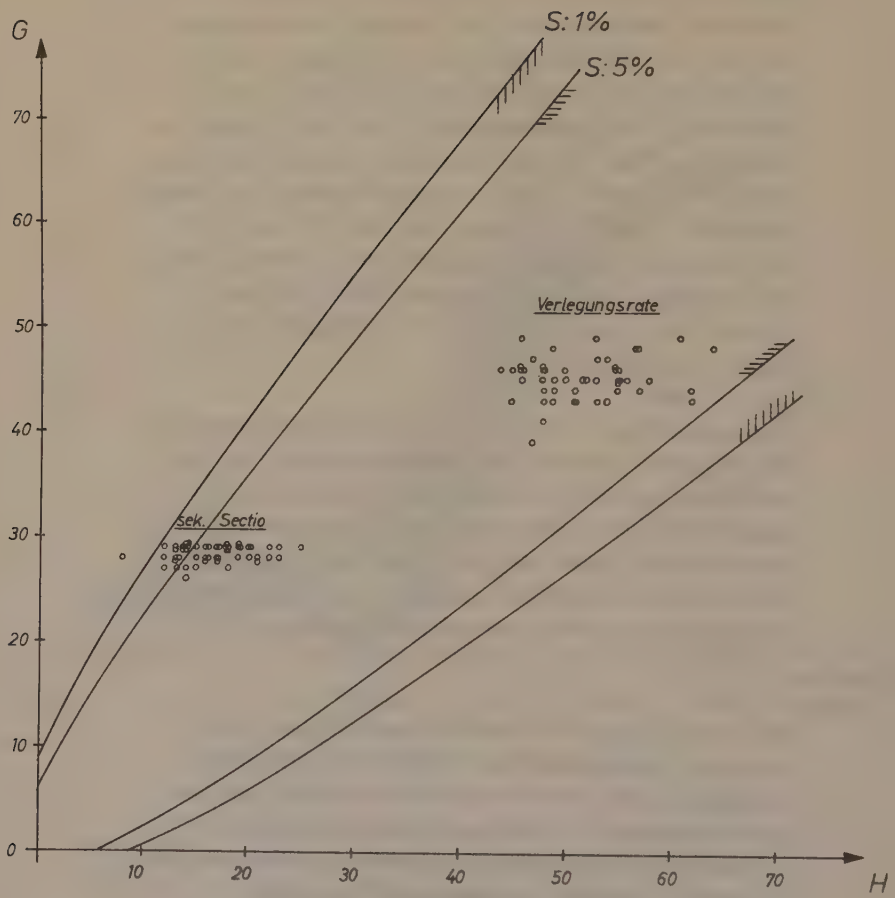


Abb. 6: Verteilung von G und H für sek. Sectio und Verlegungsrate in 50 Anpassungsverfahren.

außerhalb der Kurve ein signifikantes Ergebnis auf dem entsprechenden Signifikanzniveau kennzeichnet ($\hat{X}^2 \geq \hat{X}^2_{1,2}$). An dieser Stelle soll darauf hingewiesen werden, daß bei mehrfacher Durchführung des gleichen Tests andere Signifikanzschranken zu verwenden sind (34). Die oben angeführten Schranken gelten nur für die einmalige "isolierte" Anwendung des Tests. Die Abbildung soll in diesem Sinne nur die Streuung der erzielten Ergebnisse in Bezug auf die Signifikanzschranken des einmaligen Tests aufzeigen.

Es fällt auf, daß die Punktemengen sich nicht gleichförmig sondern parallel zur Abszisse H verteilen. Dies ist speziell bei der Punktmenge links unten der Fall, die die Ergebnisse der Abhängigkeit der sekundären Sectio vom CTG s.p. zeigt. Hier variiert die Größe H mit der Zufallszuteilung der Paare wesentlich stärker als die Größe G (maximale Differenz von G = 3, maximale Differenz von H = 17). Ein Einfluß durch die Gesamtzahl der erzielten Paare liegt nicht vor, da bei den 50 durchgeführten Anpassungsvorgängen die Zahl der erzielten Paare nur zwischen 829 und 833 schwankte.

Zur Interpretation muß man die Häufigkeiten der Prüf- und Antwortvariablen in der ungepaarten Grundgesamtheit zurückgreifen. Sie sind in der folgenden Vierfeldertafel zusammengestellt:

		sek. Sectio ja	nein	
CTG	ja	118	3973	4091
	nein	36	873	909
		154	4846	5000

Bereinigt man diese Tafel noch um die Probanden, bei denen Angaben zum Sozialstatus fehlen, ergibt sich folgendes Bild:

		sek. Sectio ja	nein	
CTG	ja	114 (A)	3885 (B)	3999
	nein	33 (C)	844 (D)	877
		147	4729	4876

Der Wert H wird bestimmt durch Kombination von Probanden der Gruppe A mit Gruppe D, der Wert G durch Kombination von Probanden der Gruppe B mit C. Da die Zahl der Paare durch die ungleiche Besetzung bei Fällen und Kontrollen in erster Linie durch die Zahl der Kontrollen limitiert wird, entstehen bei den Fällen entsprechend hohe Verluste. Diese Verluste

bewirken dann bei der geringen Besetzung der Gruppe A eine größere Variation durch die Zufallszuteilung als bei B. Dazu zwei Beispiele, die Vierfeldertafeln gleicher Anordnung wie oben, jedoch nach dem Anpassen zeigen. Einem Verlust an Fällen von 79,22% steht ein Verlust an Kontrollen von nur 5,25% gegenüber.

Beispiel 1:

		sek. Sectio ja	nein	
CTG	ja	13	818	831
	nein	29	802	831
		42	1620	1662

Beispiel 2:

		sek. Sectio ja	nein	
CTG	ja	24	807	831
	nein	29	802	831
		53	1609	1662

Bei einem durchschnittlichen Verlust an Fällen von 79% variiert dieser Verlust bei Gruppe A zwischen 79% und 88% und bei Gruppe B nur geringfügig um 79%. Die Gruppen C und D bleiben in diesen Beispielen unverändert, variieren jedoch auch insgesamt kaum. Aus der großen Variabilität von A erklärt sich auch die große Streuung von H in Abb. 6. Bei sehr ungleicher

Verteilung der Fälle und Kontrollen ist also stets mit einem Unsicherheitsfaktor durch die Zufallszuteilung von Fällen und Kontrollen zu rechnen. Dieser Unsicherheitsfaktor ist jedoch gleichzeitig ein Hinweis darauf, daß die Anpassungsvariablen nur einen geringen Einfluß auf die Abhängigkeit der Variablen "sek. Sectio" und "CTG s.p." besitzen.

4.5 Zusammenfassung.

In Hinblick auf das untersuchte Datenmaterial und unter Berücksichtigung der durch den hohen Verlust an Fällen entstandenen Unsicherheit in den X^2- Werten für die Beziehung CTG - Sectio läßt sich weder ein Einfluß der kardiotokographischen Geburtsüberwachung auf die Verlegungsrate noch auf die Beendigung der Geburt durch sekundäre Sectio ableiten. Auf ähnliche Ergebnisse wird von verschiedenen Autoren verwiesen (4, 22, 31). Im Gegensatz dazu berichtet Haverkamp (14) über eine signifikant höhere Sectiorate bei elektronisch überwachten Risikopatientinnen im Vergleich mit auskultatorisch überwachten Patientinnen. Dabei ist jedoch festzustellen, daß in dieser prospektiven randomisierten Untersuchung von Haverkamp eine selektierte Gruppe von Patientinnen mit hohem Geburtsrisiko beobachtet wurde, die sich mit unserer Untersuchungspopulation nicht vergleichen läßt. Ein direkter Vergleich beider Ergebnisse ist daher nicht möglich.

Das im Vordergrund stehende Ziel der Programmrealisierung war dasjenige, eine gültige Anpassung zu

erhalten, d.h. eine Anpassung, deren Genauigkeit durch den Anwender bestimmt wird und nicht durch eine freie Variation der Anpassungskriterien durch das Programm selbst. Im letzteren Fall wären die Kriterien, die eine Paarung bedingen, nicht mehr direkt nachvollziehbar und die Genauigkeit der Anpassung müßte in nachgeschalteten Tests überprüft werden. In den verwirklichten Programmen bleibt es dem Anwender überlassen, aus den Informationen, die ihm die einzelnen Programmteile anbieten, wie die Zahl der Paare, die Verteilung auf Qualitätsklassen und daraus resultierende Verluste oder die Verluste auf Grund der Breite des Anpassungsbereichs bei den quantitativen Variablen, Rückschlüsse auf seine Kriterien einer Anpassung zu ziehen und diese in Hinblick auf das Untersuchungsziel zu optimieren. Aus dem gleichen Grund wurde darauf verzichtet, die Ähnlichkeit zwischen Probanden durch eine gemittelte Ähnlichkeit aus allen Anpassungsvariablen zu definieren, wodurch die Eigenständigkeit jeder Variablen, ihre Gewichtung und ihre Bezüge zu anderen Variablen und die daraus resultierenden Auswirkungen verschleiert würden.

Das Problem der Zuteilung von Fall und Kontrolle bei mehreren möglichen Paarungen ist nicht allgemein, sondern nur in Bezug auf die Zielvorstellung optimal zu lösen. Hierzu böten sich prinzipiell drei Möglichkeiten an:

Erstens die Ähnlichkeit des Paares als wichtigsten Faktor in den Vordergrund zu stellen und entweder

anhand der bereits vorgegebenen Anpassungsvariablen die "ähnlichste" Paarung auszusuchen, oder über zusätzliche "fakultative" Variable die Ähnlichkeit noch enger zu definieren und dann danach die engste Anpassung auszuwählen.

Zweitens das Ziel, die Anzahl der erreichbaren Paarungen möglichst groß werden zu lassen. Dazu müßten sämtliche möglichen Paarungskombinationen gespeichert und mit Hilfe eines "Maximum-matching"-Algorithmus (8) durchsucht werden. Dieser Aufwand wird im Vergleich zum erzielbaren Effekt in den meisten Fällen nicht zu vertreten sein.

Es bleibt drittens die Zuteilung der Paare durch Zufall. Dies führt zu nur geringfügig differierenden Paarzahlen, da ihre Gesamtzahl in erster Linie durch die Definition der Ähnlichkeit bestimmt wird, wie in dem durchgeführten Beispiel gezeigt werden konnte. Zu welcher Störung der Ergebnisse dies jedoch unter der Voraussetzung sehr ungleicher Fall- und Kontrollzahlen führen kann, illustriert ebenfalls das durchgeführte Verfahren. In einem solchen Fall bietet jedoch die mehrfache Anwendung des Anpassungsverfahrens mit variablen Zufallszuteilungen dem Anwender eine Grundlage zur Abschätzung der aufgetretenen Unsicherheiten. Es bleibt weiteren Untersuchungen vorbehalten zu klären, ob und unter welchen Bedingungen das erstere dem dritten Verfahren überlegen ist.

Schließlich ist noch einzugehen auf die Auswertung und die Interpretationsmöglichkeiten der angepaßten

Untersuchungsgruppen. Da es sich bei diesem Verfahren um ein individuelles Anpassen, d.h. Anpassen von Einzelindividuen handelt, ist exakterweise nur der Test nach Mc Nemar zu verwenden (23, 2, 32), der nur die Paare mit unterschiedlichen Werten der Antwortvariablen berücksichtigt. Als hilfreich hat sich in dem realisierten Auswerteprogramm erwiesen, daß nicht nur die für den Mc Nemar-Test erforderlichen Größen G und H (Zahl der Paare mit konträren Werten der Antwortvariablen) ermittelt werden, sondern auch die ungepaarte Verteilung der Werte der Antwortvariablen, also ihre Häufigkeit bei den Fällen und Kontrollen allein. Dies erleichtert die Beurteilung von unterschiedlichen bzw. stark veränderlichen Häufigkeiten der Antwortvariablen.
Gleichfalls zur Überprüfung und Interpretation von Effekten des Anpassungsverfahrens läßt sich der Kolmogoroff-Smirnoff-Test speziell im Bereich der quantitativen Variablen verwenden. Mit ihm lassen sich Verteilungsunterschiede sowohl bei den ungepaarten als auch den gepaarten Fällen und Kontrollen feststellen oder wie im vorliegenden Anwendungsfall die Verteilung einer quantitativen Variablen (Geburtsgewicht) kontrollieren, die nicht in das Anpassungsverfahren einbezogen wurde.

Im Rahmen des Anpassungsvorgangs generierte das Verfahren aus der eingangs vorgegebenen Datei, die es nur zum Lesen verwendet und somit unverändert läßt, drei Ausgangsdateien, die verschiedene Weiterverarbeitungsmöglichkeiten eröffnen. Die erste dieser Dateien entspricht ihrem Inhalt nach exakt der Ein-

gabedatei, jedoch ist der Inhalt nach den Anpassungskriterien strukturiert. Sie ist geeignet für Untersuchungen, die spaziell das Anpassungsverfahren an sich betreffen und dazu sowohl die gepaarten als auch die ungepaarten Probanden der Untersuchungspopulation benötigen. Dies wären zum Beispiel Untersuchungen über Einflüsse und Auswirkungen des Anpassungsverfahrens auf die Zusammensetzung und die Verteilung der erzeugten Untergruppen und Restgruppen. Die beiden weiteren Ausgabedateien bieten die gepaarten Probanden in jeweils unterschiedlicher Anordnung an, so daß mit diesen Dateien weitere Auswertungen des "gematchten" Datenmaterials vorgenommen werden können.

5. Literaturverzeichnis

(1) ALTHAUSER, R.P., RUBIN, D., The computerized construction of a matched sample. American Journal of Sociology 76, 1970, S. 325-346.

(2) BILLEWICZ, W.Z., Matched samples in medical investigations. Brit. Journal prev. soc. Med. 18, 1964, S. 167-173.

(3) BILLEWICZ, W.Z., The efficiency of matched samples: an empirical investigation. Biometrics 21, 1965, S. 623-644.

(4) BOLLMANN, R., ISSEL, E.P., HALLE, H., LANGE, W., PRENZLAU, P. und HENGST, P., Einfluß der elektronischen fetalen Überwachung auf die Operationsfrequenz und den Zustand der Neugeborenen. Zbl. Gynäk. 98, 1976, S. 17-22.

(5) BROSS, I.D.J., How case-for-case matching can improve design efficiency. Am. Journal of Epidemiology 89, 1969, S. 359-363.

(6) COCHRAN, W.G., Matching in analytical studies. Am. Journal of Public Health 43, 1953, S. 684-691.

(7) DOUGLAS, J.W., 'Premature' children at primary schools. Brit. med. Journ. 1, 1960, S. 1008-1013.

(8) EVEN, S., KARIV, O., An $O(N^{2.5})$ algorithm for maximum matching in general graphs. IEEE Computer Society. 16^{th} annual Symposium on Foundations of Computer Science. Okt. 1975, S.100-112.

(9) FEINSTEIN, A.R.,Clinical biostatistics: XX. The epidemiologic trohoc, the ablative risk ratio, and 'retrospective' research. Clinical Pharmacology and Therapeutics 14, Nr. 2, 1973, S.291-307.

(1o) FISHER, L., PATIL, K., Matching unrelatedness. American Journal of Epidemiology 100, 1974, S. 347-349.

(11) FLEISS, J.L., On the asserted invariance of the odds ratio. Brit. Journal prev. soc. Med. 24, 1970, S. 45.

(12) GREENBERG, B.G., The use of analysis of covariance and balancing in analytical surveys. American Journal of Public Health 43, 1953, S. 692-699.

(13) HARDY, R.J., WHITE, C., Matching in retrospective studies. American Journal of Epidemiology 93, 1971, S. 75-76.

(14) HAVERKAMP, A.D., THOMPSON, H.E., McFEE, J.G., CETRULO, C., The evaluation of continous fetal heart rate monitoring in high-risk pregnancy. Am. J. Obstet. Gynecol., Vol. 125, Nr. 3, Juni 1976 S. 310-320.

(15) IBRAHIM, M.A., SPITZER, W.O., The case control study: the problem and the prospect. Journal of Chronic Diseases 32, 1979, S. 139-144.

(16) IMMICH, H., ECKMANN, F., Gibt das relative Risiko das Ausmaß des Krankheitsrisikos an?, Lebensversicherungsmedizin 2, 1977, S. 45-50.

(17) Mc KINLAY, S.M., The expected number of matches and its variance for matched-pair designs. Appl. Statist. 23, 1974, S. 372-383.

(18) Mc KINLAY, S.M., The effect of bias on estimators of relative risk for pair-matched and stratified samples. Journal of the American Statistical Association 70, Dec. 1975, S. 859-864.

(19) Mc KINLAY, S.M., A note on the chi-square test for pair-matched samples. Biometrics 31, 1975, 731-735.

(20) Mc KINLAY, S.M., Pair-matching - A reappraisal of a popular technique. Biometrics 33, 1977, S. 725-735.

(21) KRAUSS, B., Programmsystem "Pair matching". Technischer Bericht Nr. 9. Institut f. Medizinische Informationsverarbeitung, Statistik und Biomathematik der Ludwig-Maximilians-Universität München, 1981.

(22) LEHMANN, W.D., NEUMANN, G.K., KESSLER, K.F. und JONATHA, W.D., Operationshäufigkeit und perinatale Sterblichkeit vor und nach Einführung der fetalen Blutgasanalyse und der kontinuierlichen Überwachung der fetalen Herzfrequenz. Geburtsh. und Frauenheilk. 36, 1976, S. 247-255.

(23) MANTEL, N., HAENSZEL, W., Statistical aspects of the analysis of data from retrospective studies of disease. J. nat. Cancer Inst. 22, 1959, S. 719-748.

(24) MIETTINEN, O.S., The matched pairs design in the case of all-or-none responses. Biometrics 24, 1968, S. 339-352.

(25) MIETTINEN, O.S., Individual matching with multiple controls in the case of all-or-none responses. Biometrics 25, 1969, S. 339-355.

(26) MIETTINEN, O.S., Matching and design efficiency in retrospective studies. American Journal of Epidemiology 91, 1970, S. 111-118.

(27) MIETTINEN, O.S., Estimation of relative risk from individually matched series. Biometrics 26, 1970, S. 75-86.

(28) MIETTINEN, O.S., Confounding and effect-modification. American Journal of Epidemiology 100, 1974, S. 350-353.

(29) MIETTINEN, O.S., Simple interval-estimation of risk ratio. American Journal of Epidemiology 100 (abstract), 1974, S. 515-516.

(30) MIETTINEN, O.S., Estimability and estimation in case-referent studies. American Journal of Epidemiology 1o3, 1976, S. 226-235.

(31) PAUL, R.H., HUEY, J.R.jr. und YAEGER C.F. Clinical fetal monitoring. Its effect an cesarean section rate ans perinatal mortality: five year trends. Postgrad. Med. 61, 1977, S.160.

(32) PIKE, M.C., MORROW, R.H., Statistical analysis of patient-control studies in epidemiology. Factor under investigation on all-or-none variable. Brit. Journal prev. soc. Med. 24, 1970, S. 42-44.

(33) RUBIN, D.B., Multivariate matching methods that are equal percent bias reducing, II: Maximums on bias reduction for fixed sample sizes. Biometrics 32, 1976, S. 121-132.

(34) SACHS, L., Angewandte Statistik, Berlin, Heidelberg, New York ^{5}1978.

(35) SELBMANN, H.K., Münchner Perinatal-Studie 1975-1977. Köln-Lövenich, 1980.

(36) SMITH, A.H., KARK, J.D., CASSEL, J.C., SPEARS, G.F.S., Analysis of prospective epidemiologic studies by minimum distance case-control matching. American Journal of Epidemiology 105, 1977, S. 567-574.

(37) TAUBE, A., Matching in retrosprective studies; sampling via the dependent variable. Acta Societatis Medicorum Upsaliensis 73, 1968, S. 187-196.

(38) TAUBE, A., HEDMAN, B., On the consequences of matching in retrospective studies with special regard to the calculation of relative risks. Acta Societatis Medicorum Upsaliensis 74, 1969, S.1-16.

(39) URY, H.K., Efficiency of case-control studies with multiple controls per case: continuous or dichotomous data. Biometrics 31, 1975, S. 643-649.

(40) WORCESTER, J., Matched samples in epidemiologic studies. Biometrics 20, 1964, S. 840-848.

L e b e n s l a u f

Am 10. Dezember 1947 wurde ich, Bertram Krauß, als erstes Kind des kaufmännischen Angestellten Karl Krauß und seiner Ehefrau Charlotte in Münchberg/ Oberfranken geboren.

Dort besuchte ich von 1954 bis 1958 die Volksschule und ab 1958 das mathematisch-naturwissenschaftliche Gymnasium, an dem ich im Juli 1967 das Abitur bestand. Nach Ableistung der Wehrpflicht studierte ich ab dem Wintersemester 1969 das Fach Elektrotechnik an der Technischen Universität München. Dieses Studium beendete ich im Oktober 1974 mit der Diplom-Hauptprüfung und einer Diplomarbeit aus dem Gebiet der Nachrichtentechnik. Im Anschluß daran begann ich im Wintersemester 1974 mit dem Studium der Medizin an der Ludwig Maximilians-Universität München, das ich im Mai 1981 mit dem letzten Teil des Staatsexamens abschloß. Seit Juni 1981 bin ich als Assistenzarzt in der Klinik für Radiologie im Klinikum Großhadern beschäftigt.